(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 September 2002 (19.09.2002)

PCT

(10) International Publication Number
WO 02/072604 A2

(51) International Patent Classification[7]: C07K

(21) International Application Number: PCT/US02/06728

(22) International Filing Date: 5 March 2002 (05.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/802,640 9 March 2001 (09.03.2001) US

(71) Applicant (*for all designated States except US*): **SEQUENOM, INC.** [US/US]; 3595 John Hopkins Court, San Diego, CA 92121 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): **BRAUN, Andreas** [DE/US]; 3935 Lago Di Grata, San Diego, CA 92130 (US). **BANSAL, Aruna** [GB/GB]; 13 High Street, Landbeach, Cambs CB4 8DR (GB). **KLEYN, Patrick, W.** [NL/US]; 121 Alcott Road, Concord, MA 01742 (US).

(74) Agents: **SEIDMAN, Stephanie, L.** et al.; Heller Ehrman White & McAuliffe LLP, 4350 La Jolla Village Drive, 6th Floor, San Diego, CA 92122-1246 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)*

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/072604 A2

(54) Title: GENES AND POLYMORPHISMS ASSOCIATED WITH CARDIOVASCULAR DISEASE AND THEIR USE

(57) Abstract: Genes and polymorphisms associated with cardiovascular disease, methods that use the polymorphism to detect a predisposition to developing high cholesterol, low HDL or cardiovascular disease, to profile the response of subjects to therapeutic drugs and to develop therapeutic drugs are provided.

GENES AND POLYMORPHISMS ASSOCIATED WITH CARDIOVASCULAR DISEASE AND THEIR USE

RELATED APPLICATIONS

Benefit of priority is claimed to U.S. application Serial No. 09/802,640, entitled "GENES AND POLYMORPHISMS ASSOCIATED WITH CARDIOVASCULAR DISEASE AND THEIR USE", filed on March 9, 2001 by Andreas Braun, Aruna Bansal, and Patrick W. Kleyn. Where permitted the subject matter of this application is incorporated by reference in its entirety.

FIELD OF THE INVENTION

The field of the invention involves genes and polymorphisms of these genes that are associated with development of cardiovascular disease. Methods that use polymorphic markers for prognosticating, profiling drug response and drug discovery are provided.

BACKGROUND OF THE INVENTION

Diseases in all organisms have a genetic component, whether inherited or resulting from the body's response to environmental stresses, such as viruses and toxins. The ultimate goal of ongoing genomic research is to use this information to develop new ways to identify, treat and potentially cure these diseases. The first step has been to screen disease tissue and identify genomic changes at the level of individual samples. The identification of these "disease" markers has then fueled the development and commercialization of diagnostic tests that detect these errant genes or polymorphisms. With the increasing numbers of genetic markers, including single nucleotide polymorphisms (SNPs), microsatellites, tandem repeats, newly mapped introns and exons, the challenge to the medical and pharmaceutical communities is to identify genotypes that not only identify the disease but also follow the

progression of the disease and are predictive of an organism's response to treatment.

Polymorphisms

Polymorphisms have been known since 1901 with the identification of blood types. In the 1950's they were identified on the level of proteins using large population genetic studies. In the 1980's and 1990's many of the known protein polymorphisms were correlated with genetic loci on genomic DNA. For example, the gene dose of the apolipoprotein E type 4 allele was correlated with the risk of Alzheimer's disease in late onset families (see, *e.g.*, Corder *et al.* (1993) *Science 261*: 921-923; mutation in blood coagulation factor V was associated with resistance to activated protein C (see, *e.g.*, Bertina *et al.* (1994) *Nature 369*:64-67); resistance to HIV-1 infection has been shown in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene (see, *e.g.*, Samson *et al.* (1996) *Nature 382*:722-725); and a hypermutable tract in antigen presenting cells (APC, such as macrophages), has been identified in familial colorectal cancer in individuals of Ashkenzi jewish background (see, *e.g.*, Laken *et al.* (1997) *Nature Genet. 17*:79-83). There may be more than three million polymorphic sites in the human genome. Many have been identified, but not yet characterized or mapped or associated with a disease. Polymorphisms of the genome can lead to altered gene function, protein function or mRNA instability. To identify those polymorphisms that have clinical relevance is the goal of a world-wide scientific effort. Discovery of such polymorphisms will have a fundamental impact on the identification and development of diagnostics and drug discovery.

Single nucleotide polymorphisms (SNPs)

Much of the focus of genomics has been in the identification of SNPs, which are important for a variety of reasons. They allow indirect

testing (association of haplotypes) and direct testing (functional variants). They are the most abundant and stable genetic markers. Common diseases are best explained by common genetic alterations, and the natural variation in the human population aids in understanding disease, therapy and environmental interactions.

The organization of SNPs in the primary sequence of a gene into one of the limited number of combinations that exist as units of inheritance is termed a haplotype. Each haplotype therefore contains significantly more information than individual unorganized polymorphisms and provides an accurate measurement of the genomic variation in the two chromosomes of an individual. While it is well-established that many diseases are associated with specific variation in gene sequences and there are examples in which individual polymorphisms act as genetic markers for a particular phenotype, in other cases an individual polymorphism may be found in a variety of genomic backgrounds and therefore shows no definitive coupling between the polymorphism and the phenotype. In these instances, the observed haplotype and its frequency of occurrence in various genotypes will provide a better genetic marker for the phenotype.

Although risk factors for the development of cardiovascular disease are known, such as high serum cholesterol levels and low serum high density lipoprotein (HDL) levels, the genetic basis for the manifestation of these phenotypes remains unknown. An understanding of the genes that are responsible for controlling cholesterol and HDL levels, along with useful genetic markers and mutations in these genes that affect these phenotypes, will allow for detection of a predisposition for these risk factors and/or cardiovascular disease and the development of therapeutics to modulate such alterations.

Therefore, among the objects herein, it is an object herein to provide methods and products for detection of a predisposition for these risk factors and/or cardiovascular disease.

SUMMARY OF THE INVENTION

Provided herein are methods for using polymorphic markers to detect a predisposition to the manifestation of high serum cholesterol, low serum HDL and cardiovascular disease. The ultimate goals are the elucidation of pathological pathways, developing new diagnostic assays, determining genetic profiles for positive responses to therapeutic drugs, identifying new potential drug targets and identifying new drug candidates.

A database of twins was screened for individuals that exhibit high or low levels of serum cholesterol or HDL. Using a full genome scanning approach, SNPs present in DNA samples from these individuals were examined for alleles that associate with either high levels of cholesterol or low levels of HDL. This lead to the discovery of the association of the cytochrome C oxidase subunit VIb (COX6B) gene and the N-acetylglucosaminyl transferase component glycosylphosphatidylinositol-1 (referred to herein as GPI-1) gene with these risks factors for developing cardiovascular disease. Specifically, a previously undetermined association of an allelic variant at nucleotide 86 of the COX6B gene and high serum cholesterol levels has been discovered. In addition, it has been discovered that an allelic variant at nucleotide 2577 of the GPI-1 gene is associated with low serum HDL levels. There was no previously known association between these two genes and risk factors related to cardiovascular disease.

Methods are provided for detecting the presence or absence of at least one allelic variant associated with high cholesterol, low HDL and/or cardiovascular disease by detecting the presence or absence of at least

one allelic variant of the COX6B gene or the GPI-1 gene, individually or in combination with one or more allelic variants of other genes associated with cardiovascular disease.

Also provided are methods for indicating a predisposition to manifesting high serum cholesterol, low serum HDL and/or cardiovascular disease based on detecting the presence or absence of at least one allelic variant of the COX6B or GPI-1 genes, alone or in combination with one or more allelic variants of other genes associated with cardiovascular disease. These methods, referred to as haplotyping, are based on assaying more than one polymorphism of the COX6B and/or GPI-1 genes. One or more polymorphisms of other genes associated with cardiovascular disease may also be assayed at the same time. A collection of allelic variants of one or more genes may be more informative than a single allelic variant of any one gene. A single polymorphism of a collection of polymorphisms present in the COX6B and/or GPI-1 genes and in other genes associated with cardiovascular disease may be assayed individually or the collection may be assayed simultaneously using a multiplex assay method.

Also provided are microarrays that include a probe selected from among an oligonucleotide complementary to a polymorphic region surrounding position 86 of the sense strand of the COX6B gene coding sequence; an oligonucleotide complementary to a polymorphic region surrounding the position of the antisense strand of COX6B corresponding to position 86 of the sense strand of the COX6B gene coding sequence; an oligonucleotide complementary to a polymorphic region surrounding position 2577 of the sense strand of the GPI-1 gene; and an oligonucleotide complementary to a polymorphic region surrounding the position of the antisense strand of GPI-1 corresponding to position 2577 of the sense strand of the GPI-1 gene. Microarrays are well known and

can be made, for example, using methods set forth in U.S. Patent Nos. 5,837,832; 5,858,659; 6,043,136; 6,043,031 and 6,156,501.

Further provided are methods of using allelic variants of the COX6B or GPI-1 gene individually or together with one or more allelic variants of other genes associated with cardiovascular disease to predict a subject's response to a biologically active agent that modulates serum cholesterol, serum HDL, or a cardiovascular drug.

Also provided are methods to screen candidate biologically active agents for modulation of cholesterol, HDL or other factors associated with cardiovascular disease. These methods use cells or transgenic animals containing one or more allelic variants of the COX6B gene and/or the GPI-1 gene alone or in combination with allelic variants of one or more other genes associated with cardiovascular disease. Such animals should exhibit high cholesterol, low HDL or other known phenotypes associated with cardiovascular disease. Also, provided are methods to construct transgenic animals that are useful as models for cardiovascular disease by using one or more allelic variants of the COX6B gene and/or the GPI-1 gene alone or in combination with allelic variants of one or more other genes associated with cardiovascular disease.

Further provided are combinations of probes and primers and kits for predicting a predisposition to high serum cholesterol, low HDL levels and/or cardiovascular disease. In particular, combinations and kits contain probes or primers that are capable of hybridizing adjacent to or at polymorphic regions of the COX6B and/or GPI-1 gene. The combinations and kits can also contain probes or primers that are capable of hybridizing adjacent to or at polymorphic regions of other genes associated with cardiovascular disease. The kits also optionally contain instructions for carrying out assays, interpreting results and for aiding in diagnosing a subject as having a predisposition towards developing high serum

cholesterol, low HDL levels and/or cardiovascular disease. Combinations and kits are also provided for predicting a subject's response to a therapeutic agent directed toward modulating cholesterol, HDL, or another phenotype associated with cardiovascular disease. Such combinations and kits contain probes or primers as described above.

In particular for the methods, combinations, kits and arrays described above, the polymorphisms are SNPs. The detection or identification is of a T nucleotide at position 86 of the sense strand of the COX6B gene coding sequence or the detection or identification of an A nucleotide at the corresponding position in the antisense strand of the COX6B gene coding sequence. Also embodied is the detection or identification of an A nucleotide at position 2577 of the sense strand of the GPI-1 gene or the detection or identification of a T nucleotide at the corresponding position in the antisense strand of the GPI-1 gene. In addition to the SNPs discussed above, other polymorphisms of the COX6B and GPI-1 genes can be assayed for association with high cholesterol or low HDL, respectively, and used as disclosed above.

Other genes containing allelic variants associated with high serum cholesterol, low HDL and/or cardiovascular disease, include, but are not limited to: cholesterol ester transfer protein, plasma (CETP); apolipoprotein A-IV (APO A4); apolipoprotein A-I (APO A1); apolipoprotein E (APO E); apolipoprotein B (APO B); apolipoprotein C-III (APO C3); a gene encoding lipoprotein lipase (LPL); ATP-binding cassette transporter (ABC 1); paraoxonase 1 (PON 1);.paraoxonase 2 (PON 2); 5,10-methylenetetrahydrofolate r reductase (MTHFR); a gene encoding hepatic lipase, E-selectin, G protein beta 3 subunit and angiotensin II type 1 receptor gene.

The detection of the presence or absence of an allelic variant can use, but are not limited to, methods such as allele specific hybridization,

primer specific extension, oligonucleotide ligation assay, restriction enzyme site analysis and single-stranded conformation polymorphism analysis.

In particular, primers used in primer specific extension hybridize adjacent to nucleotide 86 of the COX6B gene or nucleotide 2577 of the GPI-1 gene or the corresponding positions on the antisense strand (numbers refer to GenBank sequences, see pages 15-17). A primer can be extended in the presence of at least one dideoxynucleotide, particularly ddG, or two dideoxynucleotides, particularly ddG and ddC. Typically, detection of extension products is by mass spectrometry. Detection of allelic variants can also involve signal moieties such as radioisotopes, enzymes, antigens, antibodies, spectrophotometric reagents, chemiluminescent reagents, fluorescent reagents and other light producing reagents.

Other probes and primers useful for the detection of allelic variants include those that hybridize at or adjacent to the SNPs described in Tables 1-3 and specifically those that include SEQ ID NOs.: 5, 10, 43, 48, 53, 58, 63, 68, 73, 78, 83, 88, 93, 98, 103, 108, 113, and 118.

DESCRIPTION OF THE DRAWINGS

Figure 1 depicts the allelic frequency and genotype for pools and individually determined samples of blood from individuals having low cholesterol levels and those with high cholesterol levels.

Figure 2 depicts the allelic frequency and genotype for pools and individually determined samples of blood from individuals having high HDL levels and those with low HDL levels.

DETAILED DESCRIPTION

A. Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which this invention belongs. All patents, patent applications and publications referred to throughout the disclosure herein are, unless noted otherwise, incorporated by reference in their entirety. In the event that there are a plurality of definitions for terms herein, those in this section prevail.

As used herein, sequencing refers to the process of determining a nucleotide sequence and can be performed using any method known to those of skill in the art. For example, if a polymorphism is identified or known, and it is desired to assess its frequency or presence in nucleic acid samples taken from the subjects that of the database, the region of interest from the samples can be isolated, such as by PCR or restriction fragments, hybridization or other suitable method known to those of skill in the art, and sequenced. For purposes herein, sequencing analysis, for example, can be effected using mass spectrometry (see, *e.g.*, U.S. Patent Nos. 5,547,835, 5,622,824, 5,851,765, and 5,928,906). Nucleic acids also can be sequenced by hybridization (see, *e.g.*, U.S. Patent Nos. 5,503,980, 5,631,134, 5,795,714) and including analysis by mass spectrometry (see, U.S. Application Serial Nos. 08/419,994 and 09/395,409). Alternatively, sequencing may be performed using other known methods, such as set forth in U.S. Patent Nos. 5,525,464; 5,695,940; 5,834,189; 5,869,242; 5,876,934; 5,908,755; 5,912,118; 5,952,174; 5,976,802; 5,981,186; 5,998,143; 6,004,744; 6,017,702; 6,018,041; 6,025,136; 6,046,005; 6,087,095; 6,117,634, 6,013,431, WO 98/30883; WO 98/56954; WO 99/09218; WO/00/58519, and the others.

As used herein, "polymorphism" refers to the coexistence of more than one form of a gene or portion thereof. A portion of a gene of which there are at least two different forms, i.e., two different nucleotide sequences, is referred to as a "polymorphic region of a gene". A polymorphic region can be a single nucleotide, the identity of which differs in different alleles. A polymorphic region also can be several nucleotides in length.

As used herein, "polymorphic gene" refers to a gene having at least one polymorphic region.

As used herein, "allele", which is used interchangeably herein with "allelic variant" refers to alternative forms of a gene or portions thereof. Alleles occupy the same locus or position on homologous chromosomes. When a subject has two identical alleles of a gene, the subject is said to be homozygous for the gene or allele. When a subject has two different alleles of a gene, the subject is said to be heterozygous for the gene. Alleles of a specific gene can differ from each other in a single nucleotide, or several nucleotides, and can include substitutions, deletions, and insertions of nucleotides. An allele of a gene also can be a form of a gene containing a mutation.

As used herein, the term "subject" refers to mammals and in particular human beings.

As used herein, the term "gene" or "recombinant gene" refers to a nucleic acid molecule comprising an open reading frame and including at least one exon and (optionally) at least one intron sequence. A gene can be either RNA or DNA. Genes may include regions preceding and following the coding region (leader and trailer).

As used herein, "intron" refers to a DNA sequence present in a given gene that is spliced out during mRNA maturation.

As used herein, the term "coding sequence" refers to that portion of a gene that encodes an amino acid sequence of a protein.

As used herein, the term "sense strand" refers to that strand of a double-stranded nucleic acid molecule that encodes the sequence of the mRNA that encodes the amino acid sequence encoded by the double-stranded nucleic acid molecule.

As used herein, the term "antisense strand" refers to that strand of a double-stranded nucleic acid molecule that is the complement of the sequence of the mRNA that encodes the amino acid sequence encoded by the double-stranded nucleic acid molecule.

As used herein, a DNA or nucleic acid homolog refers to a nucleic acid that includes a preselected conserved nucleotide sequence. By the term "substantially homologous" is meant having at least 80%, preferably at least 90%, most preferably at least 95% homology therewith or a less percentage of homology or identity and conserved biological activity or function.

Regarding hybridization, as used herein, stringency conditions to achieve specific hybridization refer to the washing conditions for removing the non-specific probes or primers and conditions that are equivalent to either high, medium, or low stringency as described below:

1) high stringency: 0.1 x SSPE, 0.1% SDS, 65°C
2) medium stringency: 0.2 x SSPE, 0.1% SDS, 50°C
3) low stringency: 1.0 x SSPE, 0.1% SDS, 50°C.

It is understood that equivalent stringencies may be achieved using alternative buffers, salts and temperatures.

As used herein, "heterologous DNA" is DNA that encodes RNA and proteins that are not normally produced *in vivo* by the cell in which it is expressed or that mediates or encodes mediators that alter expression of endogenous DNA by affecting transcription, translation, or other

regulatable biochemical processes or is not present in the exact orientation or position as the counterpart DNA in a wildtype cell. Heterologous DNA may also be referred to as foreign DNA. Any DNA that one of skill in the art would recognize or consider as heterologous or foreign to the cell in which is expressed is herein encompassed by heterologous DNA. Examples of heterologous DNA include, but are not limited to, DNA that encodes traceable marker proteins, such as a protein that confers drug resistance, DNA that encodes therapeutically effective substances, such as anti-cancer agents, enzymes and hormones, and DNA that encodes other types of proteins, such as antibodies. Antibodies that are encoded by heterologous DNA may be secreted or expressed on the surface of the cell in which the heterologous DNA has been introduced.

As used herein, a "promoter region" refers to the portion of DNA of a gene that controls transcription of the DNA to which it is operatively linked. The promoter region includes specific sequences of DNA that are sufficient for RNA polymerase recognition, binding and transcription initiation. This portion of the promoter region is referred to as the promoter. In addition, the promoter region includes sequences that modulate this recognition, binding and transcription initiation activity of the RNA polymerase. These sequences may be *cis* acting or may be responsive to *trans* acting factors. Promoters, depending upon the nature of the regulation, may be constitutive or regulated.

As used herein, the phrase “operatively linked” generally means the sequences or segments have been covalently joined into one piece of DNA, whether in single or double stranded form, whereby control or regulatory sequences on one segment control or permit expression or replication or other such control of other segments. The two segments are not necessarily contiguous. For gene expression a DNA sequence and

a regulatory sequence(s) are connected in such a way to control or permit gene expression when the appropriate molecular, e.g., transcriptional activator proteins, are bound to the regulatory sequence(s).

As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One exemplary type of vector is an episome, i.e., a nucleic acid capable of extra-chromosomal replication. Exemplary vectors include those capable of autonomous replication and/or expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively linked are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" that refer generally to circular double stranded DNA loops which, in their vector form are not bound to the chromosome. "Plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. Also included are other forms of expression vectors that serve equivalent functions and that become known in the art subsequently hereto.

As used herein, "indicating" or "determining" means that the presence or absence of an allelic variant may be one of many factors that are considered when a subject's predisposition to a disease or disorder is evaluated. Thus a predisposition to a disease or disorder is not necessarily conclusively determined by only ascertaining the presence or absence of one or more allelic variants, but the presence of one of more of such variants is among an number of factors considered.

As used herein, "predisposition to develop a disease or disorder" means that a subject having a particular genotype and/or haplotype has a higher likelihood than one not having such a genotype and/or haplotype for developing a particular disease or disorder.

As used herein, "transgenic animal" refers to any animal, generally a non-human animal, *e.g.* a mammal, bird or an amphibian, in which one or more of the cells of the animal contain heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. This molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA. In the typical transgenic animals described herein, the transgene causes cells to express a recombinant form of a protein. However, transgenic animals in which the recombinant gene is silent are also contemplated, as for example, using the FLP or CRE recombinase dependent constructs. Moreover, "transgenic animal" also includes those recombinant animals in which gene disruption of one or more genes is caused by human intervention, including both recombination and antisense techniques.

As used herein, "transgene" describes genetic material that has been or is about to be artificially inserted into the genome of a mammalian cell, particularly a mammalian cell of a living animal. The transgene is used to transform a cell, meaning that a permanent or transient genetic change, typically a permanent genetic change, is induced in a cell following incorporation of exogenous DNA. A permanent genetic change is generally achieved by introduction of the DNA into the genome of the cell. Vectors for stable integration include, but are not limited to, plasmids, retroviruses and other animal viruses and YACS. Of interest are

transgenic mammals, including, but are not limited to, cows, pigs, goats, horses and others, and particularly rodents, including rats and mice.

As used herein, "associated" refers to coincidence with the development or manifestation of a disease, condition or phenotype. Association may be due to, but is not limited to, genes responsible for housekeeping functions, those that are part of a pathway that is involved in a specific disease, condition or phenotype and those that indirectly contribute to the manifestation of a disease, condition or phenotype.

As used herein, "high serum cholesterol" refers to a level of serum cholesterol that is greater than that considered to be in the normal range for a given age in a population, e.g., about 5.25 mmoles/L or greater, *i.e.*, approximately one standard deviation or more away from the age-adjusted mean.

As used herein, "low serum HDL" refers to a level of serum HDL that is less than that considered to be in the normal range for a given age in a population, e.g. about 1.11 mmoles/L or less, *i.e.*, approximately one standard deviation or more away from the age-adjusted mean.

As used herein, "cardiovascular disease" refers to any manifestation of or predisposition to cardiovascular disease including, but not limited to, coronary artery disease and myocardial infarction. Included in predisposition is the manifestation of risks factors such as high serum cholesterol levels and low serum HDL levels.

As used herein, "target nucleic acid" refers to a nucleic acid molecule that contains all or a portion of a polymorphic region of a gene of interest.

As used herein, "signal moiety" refers to any moiety that allows for the detection of a nucleic acid molecule. Included are moieties covalently attached to nucleic acids and those that are not.

As used herein, "biologically active agent that modulates serum cholesterol" refers to any drug, including, but are not limited to, small molecule, nucleic acid (sense and antisense), protein, peptide, lipid, carbohydrate and combinations thereof, that exhibits some effect directly or indirectly on the cholesterol measured in a subject's serum.

As used herein, "biologically active agent that modulates serum HDL" refers to any drug, such as, but are not limited to, small molecule, nucleic acid (sense and antisense), protein, peptide, lipid, carbohydrate and combinations thereof that exhibits some effect directly or indirectly on the HDL measured in a subject's serum.

As used herein, "expression and/or activity" refers to the level of transcription or translation of the COX6B or GPI-1 gene, mRNA stability, protein stability or biological activity.

As used herein, "cardiovascular drug" refers to a drug used to treat cardiovascular disease or a risk factor for the disease, either prophylactically or after a risk factor or disease condition has developed. Cardiovascular drugs include those drugs used to lower serum cholesterol and those used to alter the level of serum HDL.

As used herein, "combining" refers to contacting the biologically active agent with a cell or animal such that the agent is introduced into the cell or animal. For a cell any method that results in an agent traversing the plasma membrane is useful. For an animal any of the standard routes of administration of an agent, *e.g.* oral, rectal, transmucosal, intestinal, intravenous, intraperitoneal, intraventricular, subcutaneous, intramuscular and other routes can be used.

As used herein,"positive response" refers to improving or ameliorating at least one symptom or detectable characteristic of a disease or condition, *e.g.*, lowering serum cholesterol levels or raising serum HDL levels.

As used herein, "biological sample" refers to any cell type or tissue of a subject from which nucleic acid, particularly DNA, can be obtained.

As used herein, "array" refers to a collection of three or more items, such a collection of immobilized nucleic acid probes arranged on a solid substrate, such as silica, polymeric materials, glass and other suitable support materials known to those of skill in the art.

As used herein, a composition refers to any mixture. It may be a solution, a suspension, liquid, powder, a paste, aqueous, non-aqueous or any combination thereof.

As used herein, a combination refers to any association between two or among more items.

As used herein, "kit" refers to a package that contains a combination, such as one or more primers or probes used to amplify or detect polymorphic regions of genes associated with cardiovascular disease, optionally including instructions and/or reagents for their use.

As used herein "specifically hybridizes" refers to hybridization of a probe or primer only to a target sequence preferentially to a non-target sequence. Those of skill in the art are familiar with parameters that affect hybridization; such as temperature, probe or primer length and composition, buffer composition and salt concentration and can readily adjust these parameters to achieve specific hybridization of a nucleic acid to a target sequence.

As used herein "nucleic acid" refers to polynucleotides such as deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). The term should also be understood to include, as equivalents, derivatives, variants and analogs of either RNA or DNA made from nucleotide analogs, single (sense or antisense) and double-stranded polynucleotides. Deoxyribonucleotides include deoxyadenosine, deoxycytidine, deoxyguanosine and deoxythymidine. For RNA, the uracil base is uridine.

As used herein, "mass spectrometry" encompassses any suitable mass spectrometric format known to those of skill in the art. Such formats include, but are not limited to, Matrix-Assisted Laser Desorption/Ionization, Time-of-Flight (MALDI-TOF), Electrospray (ES), IR-MALDI (see, e.g., published International PCT Application No. 99/57318 and U.S. Patent No. 5,118,937) Ion Cyclotron Resonance (ICR), Fourier Transform and combinations thereof. MALDI, particular UV and IR, are among exemplary formats.

As used herein, the GPI-1 gene is generically used to include the human GPI-1 gene and its homologs from rat, mouse, guinea pig, mouse and other mammalian species. As described below, the GPI-1 gene refers to a component of the GlcNAc transferase activity complex that functions in the biosynthesis of glycosylphosphatidylinositol (GPI) anchor. Four mammalian gene products (PIG-A, PIG-H, PIG-C and GPI-1) form a protein complex that is responsible for the transferase enzyme activity in the biosynthesis reaction . PIG-A, PIG-H, PIG-C are required for the first step in GPI anchor biosynthesis; GPI-1 is not. Stabilization of the enzyme complex, rather than participation in GlcNAc transfer, has been suggested as a possible role for GPI-1 (Watanabe *et al.* EMBO 17:877, 1998).

B. Cytochrome c oxidase VIb gene

Cytochrome c oxidase (COX) is a mitochondrial enzyme complex integrated in the inner membrane. It transfers electrons from cytochrome to molecular oxygen in the terminal reaction of the respiratory chain in eukaryotic cells. COX contains of three large subunits encoded by the mitochondrial genome and 10 other subunits, encoded by nuclear genes. The three subunits encoded by mitochondrial genome are responsible for the catalytic activity. The cytochrome c oxidase subunit VIb (COX6B) is one of the nuclear gene products. The function of the nuclear encoded subunits is unknown. One proposed role is in the regulation of catalytic

activity; specifically the rate of electron transport and stoichiometry of proton pumping. Other proposed roles are not directly related to electron transport and include energy-dependent calcium uptake and protein import by the mitochondrion. Proteolytic removal of subunits VIa and VIb has been associated with loss of calcium transport in reconstituted vesicles. Steady-state levels of the COX6B transcript are different in different tissues (Taanman *et al.*, Gene (1990), 93:285). The COX6B gene is includes the human COX6B gene and its homologs from rat, mouse, guinea pig, and any species that has a homologous gene.

Several single nucleotide polymorphism have been identified in the human COX6B gene. One of these is located at position 86 and is a C to T transversion that is manifested as a silent mutation in the coding region, ACC to ACT (threonine to threonine)(SEQ ID NO.: 2). Although this is a silent mutation at the amino acid level, it may represent an alteration that changes codon usage, or it may effect mRNA stability or it may be in linkage disequilibrium with a non-silent change. Other known single nucleotide polymorphisms of the COX6B gene include, but are not limited to, those listed in Table 1.

TABLE 1

Gene	GenBank Accession No.	SNP	SNP Location
COX6B (SEQ ID NO.: 1)	**NM_001863**	C/T	86
		A/G	60
		A/T	324
		A/T	123

Based on methods disclosed herein and those used in the art, one of skill would be able to use all the SNPs described and find additional polymorphic regions of the COX6B gene to determine whether allelic variants of these regions are associated with high cholesterol levels and cardiovascular disease.

C. GPI-1 Gene

Glycosylphosphatidylinositol (GPI) functions to anchor various eukaryotic proteins to membranes and is essential for their surface expression. Thus, a defect in GPI anchor synthesis affects various functions of cell, tissues and organs. Biosynthesis of glycosyl-phosphatidylinositol (GPI) is initiated by the transfer of N-acetylglucosamine (GlcNAc) from UDP-GlcNac to phosphatidylinositol (PI) and is catalyzed by a GlcNAc transferase, GPI-GlcNAc transferase (GPI-GnT). Four mammalian gene products form a protein complex that is responsible for this enzyme activity (PIG-A, PIG-H, PIG-C and GPI-1). PIG-A, PIG-H, PIG-C are required for the first step in GPI anchor biosynthesis; GPI-1 is not. Stabilization of the enzyme complex, rather than participation in GlcNAc transfer, has been suggested as a possible role for GPI-1 (Watanabe *et al.* EMBO 17:877, 1998).

A polymorphism has been identified at position 2577 of the human GPI-1 gene. This is a G to A transversion. This SNP is located in the 3' untranslated region of the mRNA, and does not affect protein structure, but may affect mRNA stability or may be in linkage disequilibrium with a non-silent change. Other known single nucleotide polymorphisms of the GPI-1 gene include, but are not limited to, those listed in Table 2.

TABLE 2

Gene	GenBank Accession No.	SNP	SNP Location
GPI-1 (SEQ ID NOS.: 6, 7)	NM_004204	C/T	2829
		A/G	2577
		C/T	2519
		C/T	2289
		C/T	1938
		C/G	1563
		A/G/C/T	2664
		A/G	2656
		A/C/T	2167
		G/C/A	2166

Based on methods disclosed herein and those used in the art, one of skill would be able to use all the described SNPs and find additional polymorphic regions of the GPI-1 gene to determine whether allelic variants of these regions are associated with low levels of HDL and cardiovascular disease.

D. Other genes and polymorphism associated with cardiovascular disease

Many other genes and polymorphisms contained within them have been associated with risks factors for cardiovascular disease (aberrations in lipid metabolism; specifically high levels of serum cholesterol and low levels of HDL and other such indicators) and/or the clinical phenotypes of atherosclerosis and cardiovascular disease. Table 3 presents a list of some of these genes and some associated polymorphisms (SNPs): cholesterol ester transfer protein, plasma (CETP); apolipoprotein A-IV (APO A4); apolipoprotein A-I (APO A1); apolipoprotein E (APO E); apolipoprotein B (APO B); apolipoprotein C-III (APO C3); a gene encoding lipoprotein lipase (LPL); ATP-binding cassette transporter (ABC 1); paraoxonase 1 (PON 1); paraoxonase 2 (PON 2); 5,10-methylenetetrahydrofolate r reductase (MTHFR); a gene encoding hepatic

lipase (LIPC); E-selectin; G protein beta 3 subunit and angiotensin II type 1 receptor gene. The SNP locations are based on the GenBank sequence. Table 3 is not meant to be exhaustive, as one of skill in the art based on the disclosure would be able to readily use other known polymorphisms in these and other genes, new polymorphisms discovered in previously identified genes and newly identified genes and polymorphisms in the methods and compositions disclosed herein.

TABLE 3

Gene	GenBank Accession No.	SNP	SNP Location
CETP (SEQ ID NOS.: 11, 12)	**NM_000078**	C/A	991
		C/T	196
		A/G	1586
		A/G	1394
		A/G	1439
		C/G	1297
		C/T	766
		G/A	1131
		G/A	1696
LPL (SEQ ID NOS.: 13, 14)	**NM_000237**	A/G	1127
		A/C	3447
		C/T	1973
		C/T	3343
		G/A	2851
		C/T	3272
		A/T	2428
		T/C	2743
		G/A	1453
		C/A	3449
		G/A	1282
		G/A	579
		A/C	1338
		A/G/T/C	2416-2426
		A/G	2427
		C/T	1302
		G/A	609

TABLE 3

		G/C	1595
		G/A	1309
		C/T	2454
		C/T	2988
		G/A	280
		G/A	1036
APO A4 (SEQ ID NOS.: 15, 16)	**NM_000482**	G/T	1122
		G/C	1033
		G/A	1002
		C/T	960
		C/T	894
		G/A	554
		G/A	950
		T/C	336
		G/A	334
		C/T	330
		A/G	201
		A/G	16
		A/T	1213
APO E (SEQ ID NOS.: 17, 18) (mRNA)	**NM_000041**	C/T	448
		G/A	448
		C/T	586
		C/T	197
		C/T	540
Hepatic Lipase (SEQ ID NOS.: 19, 20)	**NM_000236**	C/G	680
		G/A	1374
		G/A	701
		C/A	1492
		A/G	648
		G/C	729
		G/A	340
		G/T	522
PON 1 (SEQ ID NOS.: 21, 22)	**NM_000446**	A/T	172
		A/G	584
		G/C	190

TABLE 3

PON 2 (SEQ ID NOS.: 23, 24)	**XM_004947**	C/G	475
		C/G	964
APO C3 (SEQ ID NOS.: 25, 26)	**NM_000040**	C/T	148
		T/A	471
		G/C	386
		G/T	417
		T/A	495
ABC 1 (SEQ ID NOS.: 27, 28)	**XM_005567**	G/A	8591
APO A1 (SEQ ID NOS.: 29, 30)	**NM_000039**	C/G	770
		G/A	656
		C/G	589
		C/G	414
		A/T	430
		C/T	708
		C/T	221
		T/G	223
		C/T	597
		A/G	340
		G/C	690
APO B (SEQ ID NOS.: 31, 32)	**NM_000384**	A/G/C/T	13141
		A/G/C/T	12669
		C/T	11323
		G/C	10422
		A/C	10408
		C/G	10083
		C/T	7064
		C/T	6666
		C/T	1980
		C/G	5751
		C/T	7673
		C/A/G/T	8344
		G/C/T/A	4393
		A/C/T/G	5894
		A/T	12019
		C/T	11973

TABLE 3

		G/C/T/A	7065
		C/G	947
		C/G	7331
		A/G	7221
		G/C	6402
		G/C	3780
		C/G	1661
		A/T	8167
		C/A	8126
		C/T	421
		C/T	1981
		G/A	12510
		G/C	12937
APO B (con't)		G/A	11042
		C/T	2834
		A/G	5869
		A/G	11962
		C/G	4439
		G/A	7824
		G/A	13569
		G/A	9489
		G/A	2325
		G/A	10259
		C/G	14
MTHFR (SEQ ID NOS.: 33, 34)	**NM_005957**	G/A	5442
		A/G	5113
		A/G	5113
		A/G	5110
		A/G	5102
		A/C/T	5097
		A/C/T	5097
		C/T	5079
		C/T	5079
		T/C	5071
		T/C	5071
		T/C	5051

TABLE 3

		G/A	5012
		C/A	5000
		A/G	4998
		A/G	4994
		A/G	4994
		A/G	4994
		C/T	4991
		C/T	4991
		C/T	4991
		A/G	4986
		A/G	4986
		A/G	4986
		C/T	4985
		T/A	4982
		T/G	4981
		T/C	4981
		T/C	4981
MTHFR (con't)		G/C/A	4967
		G/A	4963
		A/G	4962
		G/C/T	4962
		A/C/G/T	4961
		A/C/T	4961
		A/C	4961
		A/C	4961
		A/C/T	4960
		T/C	4938
		T/C	4937
		T/C	4933
		G/C/T	4933
		C/T	4929
		C/T	4929
		T/A/G	4929
		A/G	4928
		G/C	4928
		C/G	4927

TABLE 3

		G/A	4923
		C/T	4919
		A/T/G	4913
		C/T	4912
		A/T	4903
		C/T	4902
		A/G	4900
		G/A	4898
		G/T	4898
		C/T	4897
		G/T	4894
		T/C/G	4836
		C/T	3862
		C/T	4922
		C/T	4959
		T/C	4981
		A/G	4994
		A/G	5044
		T/C	5051
		G/C	5066
		C/T	5079
MTHFR (con't)		C/A/G	5085
		C/T	5092
		A/G	5103
		A/G	5113
		C/T	1021
E-Selectin (SEQ ID NOS.: 35, 36)	**NM_000450**	G/A	3484
		G/A	3093
		T/G	2939
		T/C	2902
		C/T	1937
		C/T	1916
		C/T	1839
		C/T	1805
		C/T	1518
		G/C	1377

TABLE 3

		C/T	1376
		G/A	999
		T/C	857
		A/C	561
		C/G	506
		A/G	392
		G/T	98
G protein β3 subunit (SEQ ID NOS.: 37, 38)	**NM_002075**	C/T	1828
		C/T	1546
		G/T	1431
		G/A	1231
		C/T	1230
Angiotensin II type 1 receptor gene (SEQ ID NOS.: 39, 40)	**NM_00686**	G/A	1453
		C/G	968
		G/C	966
		T/C	941
		G/A	894
		T/C	659

Assays to identify the nucleotide present at the polymorphic site include those described herein and all others known to those who practice the art.

For some of the SNPs described above, there are provided a description of the MassEXTEND™ reaction components that can be used to determine the allelic variant that is present. Included are the forward and reverse primers used for amplification. Also included are the MassEXTEND™ primer used in the primer extension reaction and the extended MassEXTEND™ primers for each allele. MassEXTEND™ reactions are carried out and the products analyzed as described in Examples 2 and 3.

CETP

Position 991 (C/A)

PCR primers:

Forward:	ACTGCCTGATAACCATGCTG (SEQ ID NO.: 41)
Reverse:	ATACTTACACACCAGGAGGG (SEQ ID NO.: 42)
MassEXTEND™ Primer:	ATGCCTGCTCCAAAGGCAC (SEQ ID NO.: 43)
Primer Mass:	5757.8
Extended Primer-Allele C:	ATGCCTGCTCCAAAGGCACC (SEQ ID NO.: 44)
Extended Primer Mass:	6030.9
Extended Primer-Allele A:	ATGCCTGCTCCAAAGGCACAT (SEQ ID NO.: 45)
Extended Primer Mass:	6359.2

Position 196 (C/T)

PCR primers:

Forward:	TACTTCTGGTTCTCTGAGCG (SEQ ID NO.: 46)
Reverse:	ACTCACCTTGAACTCGTCTC (SEQ ID NO.: 47)
MassEXTEND™ Primer:	TGGTTCTCTGAGCGAGTCTT (SEQ ID NO.: 48)
Primer Mass:	6130
Extended Primer-Allele C:	TGGTTCTCTGAGCGAGTCTTC

	(SEQ ID NO.: 49)
Extended Primer Mass:	6707.4
Extended Primer-Allele T:	TGGTTCTCTGAGCGAGTCTTTC (SEQ ID NO.: 50)
Extended Primer Mass:	6333.1

Position 1586 (A/G)

PCR primers:

Forward:	TGCAGATGGACTTTGGCTTC (SEQ ID NO.: 51)
Reverse:	TGCTTGCCTTCTGCTACAAG (SEQ ID NO.: 52)
MassEXTEND™ Primer:	CTTCCCTGAGCACCTGCTG (SEQ ID NO.: 53)
Primer Mass:	5715.7
Extended Primer-Allele G:	CTTCCCTGAGCACCTGCTGGT (SEQ ID NO.: 54)
Extended Primer Mass:	6333.1
Extended Primer-Allele A:	CTTCCCTGAGCACCTGCTGA (SEQ ID NO.: 55)
Extended Primer Mass:	6012.9

<u>APOA4</u>

Position 1122 (G/T)

PCR primers:

Forward:	AACAGCTCAGGACGAAACTG (SEQ ID NO.: 56)

Reverse:	AGAAGGAGTTGACCTTGTCC (SEQ ID NO.: 57)
MassEXTEND™ Primer:	GGAAGCTCAAGTGGCCTTC (SEQ ID NO.: 58)
Primer Mass:	5828.8
Extended Primer-Allele G:	GGAAGCTCAAGTGGCCTTCC (SEQ ID NO.: 59)
Extended Primer Mass:	6102.0
Extended Primer-Allele T:	GGAAGCTCAAGTGGCCTTCAAC (SEQ ID NO.: 60)
Extended Primer Mass:	6728.4

Position 1033 (G/C)

PCR primers:

Forward:	AAGTCACTGGCAGAGCTGG (SEQ ID NO.: 61)
Reverse:	GCACCAGGGCTTTGTTGAAG (SEQ ID NO.: 62)
MassEXTEND™ Primer:	TTTTCCCCGTAGGGCTCCA (SEQ ID NO.: 63)
Primer Mass:	5730.7
Extended Primer-Allele G:	TTTTCCCCGTAGGGCTCCAC (SEQ ID NO.: 64)
Extended Primer Mass:	6003.9
Extended Primer-Allele C:	TTTTCCCCGTAGGGCTCCAGC (SEQ ID NO.: 65)
Extended Primer Mass:	6333.1

Position 1002 (G/A)

PCR primers:

Forward:	TGCAGAAGTCACTGGCAGAG (SEQ ID NO.: 66)
Reverse:	GTTGAAGTTTTCCCCGTAGG (SEQ ID NO.: 67)
MassEXTEND™ Primer:	ACTCCTCCACCTGCTGGTC (SEQ ID NO.: 68)
Primer Mass:	5675.7
Extended Primer-Allele G:	ACTCCTCCACCTGCTGGTCC (SEQ ID NO.: 69)
Extended Primer Mass:	5948.9
Extended Primer-Allele A:	ACTCCTCCACCTGCTGGTCTA (SEQ ID NO.: 70)
Extended Primer Mass:	6277.1

Position 960 (C/T)

PCR primers:

Forward:	AGGACGTGCGTGGCAACCTG (SEQ ID NO.: 71)
Reverse:	AGCTCTGCCAGTGACTTCTG (SEQ ID NO.: 72)
MassEXTEND™ Primer:	GTGACTTCTGCAGCCCCTC (SEQ ID NO.: 73)
Primer Mass:	5715.7
Extended Primer-Allele T:	GTGACTTCTGCAGCCCCTCA (SEQ ID NO.: 74)

Extended Primer Mass:	6012.9
Extended Primer-Allele C:	GTGACTTCTGCAGCCCCTCGGT (SEQ ID NO.: 75)
Extended Primer Mass:	6662.3

Position 894 (C/T)

PCR primers:

Forward:	CCTGACCTTCCAGATGAAG (SEQ ID NO.: 76)
Reverse:	TCAGGTTGCCACGCACGTC (SEQ ID NO.: 77)
MassEXTEND™ Primer:	CAGGATCTCGGCCAGTGC (SEQ ID NO.: 78)
Primer Mass:	5500.6
Extended Primer-Allele C:	CAGGATCTCGGCCAGTGCC (SEQ ID NO.: 79)
Extended Primer Mass:	5773.8
Extended Primer-Allele T:	CAGGATCTCGGCCAGTGCTG (SEQ ID NO.: 80)
Extended Primer Mass:	6118.0

Position 554 (G/A)
PCR primers:

Forward:	ACCTGCGAGAGCTTCAGCAG (SEQ ID NO.: 81)
Reverse:	TCTCCATGCGCTGTGCGTAG (SEQ ID NO.: 82)
MassEXTEND™ Primer:	AGCTGCGCACCCAGGTCA (SEQ ID NO.: 83)

Primer Mass:	5469.6
Extended Primer-Allele A:	AGCTGCGCACCCAGGTCAA (SEQ ID NO.: 84)
Extended Primer Mass:	5766.8
Extended Primer-Allele G:	AGCTGCGCACCCAGGTCAGC (SEQ ID NO.: 85)
Extended Primer Mass:	6072.0

<u>APOE</u>

Position 448 (C/T)
PCR primers:

Forward:	TGTCCAAGGAGCTGCAGGC (SEQ ID NO.: 86)
Reverse:	CTTACGCAGCTTGCGCAGGT (SEQ ID NO.: 87)
MassEXTEND™ Primer:	GCGGACATGGAGGACGTG (SEQ ID NO.: 88)
Primer Mass:	5629.7
Extended Primer-Allele C:	GCGGACATGGAGGACGTGC (SEQ ID NO.: 89)
Extended Primer Mass:	5902.8
Extended Primer-Allele T:	GCGGACATGGAGGACGTGTG (SEQ ID NO.: 90)
Extended Primer Mass:	6247.1

LPL

Position 1127 (A/G)
PCR primers:

Forward:	GTTGTAGAAAGAACCGCTGC (SEQ ID NO.: 91)
Reverse:	GAGAACGAGTCTTCAGGTAC (SEQ ID NO.: 92)
MassEXTEND™ Primer:	ACAATCTGGGCTATGAGATCA (SEQ ID NO.: 93)
Primer Mass:	6454.2
Extended Primer-Allele A:	ACAATCTGGGCTATGAGATCAA (SEQ ID NO.: 94)
Extended Primer Mass:	6751.4
Extended Primer-Allele G:	ACAATCTGGGCTATGAGATCAGT (SEQ ID NO.: 95)
Extended Primer Mass:	7071.6

Position 3447 (A/C)
PCR primers:

Forward:	CACTCTACACTGCATGTCTC (SEQ ID NO.: 96)
Reverse:	ACCCTTCTGAAAAGGAGAGG (SEQ ID NO.: 97)
MassEXTEND™ Primer:	GAGGAGAGACAAGGCAGATA (SEQ ID NO.: 98)
Primer Mass:	6273.1
Extended Primer-Allele A:	GAGGAGAGACAAGGCAGATAT (SEQ ID NO.: 99)

Extended Primer Mass:	6561.3
Extended Primer-Allele C:	GAGGAGAGACAAGGCAGATAGT (SEQ ID NO.: 100)
Extended Primer Mass:	6890.5

Position 1973 (C/T)
PCR primers:

Forward:	AAAGGTTCAGTTGCTGCTGC (SEQ ID NO.: 101)
Reverse:	GCTGGGGAAGGTCTAATAAC (SEQ ID NO.: 102)
MassEXTEND™ Primer:	GTTGCTGCTGCCTCGAATC (SEQ ID NO.: 103)
Primer Mass:	5770.7
Extended Primer-Allele C:	GTTGCTGCTGCCTCGAATCC (SEQ ID NO.: 104)
Extended Primer Mass:	6043.9
Extended Primer-Allele T:	GTTGCTGCTGCCTCGAATCTG (SEQ ID NO.: 105)
Extended Primer Mass:	6388.2

<u>LIPC</u>

Position 680 (C/G)
PCR primers:

Forward:	CGTCTTTCTCCAGATGATGC (SEQ ID NO.: 106)
Reverse:	AGTGTCCTATGGGCTGTTTG (SEQ ID NO.: 107)
MassEXTEND™ Primer:	GGATGCCATTCATACCTTTAC

	(SEQ ID NO.: 108)
Primer Mass:	6556.1
Extended Primer-Allele C:	GGATGCCATTCATACCTTTACC (SEQ ID NO.: 109)
Extended Primer Mass:	6629.3
Extended Primer-Allele G:	GGATGCCATTCATACCTTTACGC (SEQ ID NO.: 110)
Extended Primer Mass:	6958.5

Position 1374 (G/A)
PCR primers:

Forward:	TGGGAAAACAGTGCAGTGTG (SEQ ID NO.: 111)
Reverse:	TGATCGTCTTCAGAACGAGG (SEQ ID NO.: 112)
MassEXTEND™ Primer:	CCAGACCATCATCCCATGGA (SEQ ID NO.: 113)
Primer Mass:	6030.9
Extended Primer-Allele A:	CCAGACCATCATCCCATGGAA (SEQ ID NO.: 114)
Extended Primer Mass:	6328.1
Extended Primer-Allele G:	CCAGACCATCATCCCATGGAGC (SEQ ID NO.: 115)
Extended Primer Mass:	6633.3

Position 701 (G/A)
PCR primers:

Forward:	CAGCAATCGTCTTTCTCCAG (SEQ ID NO.: 116)
Reverse:	TCCTATGGGCTGTTTGATGC (SEQ ID NO.: 117)
MassEXTEND™ Primer:	GTCTTTCTCCAGATGATGCCA (SEQ ID NO.: 118)
Primer Mass:	6372.2
Extended Primer-Allele A:	GTCTTTCTCCAGATGATGCCAA (SEQ ID NO.: 119)
Extended Primer Mass:	6669.4
Extended Primer-Allele G:	GTCTTTCTCCAGATGATGCCAGT (SEQ ID NO.: 120)
Extended Primer Mass:	6989.6

E. Databases

Databases for determining an association between polymorphic regions of genes and intermediate and clinical phenotypes, contain biological samples (*e.g.*, blood) that provide a source of nucleic acid and clinical data covering diseases (*e.g.*, age, sex, ethnicity medical history and family medical history) from both individuals exhibiting the phenotype (intermediate phenotype (risk factor) or clinical phenotype (disease)) and those who do not. These databases include human population groups such as twins, diverse affected families, isolated founder populations and drug trial subjects. The quality and consistency of the clinical resources are of primary importance.

F. Association Studies

The examples set forth below used an extreme trait analysis to discover an association between an allelic variant of the COX6B gene and high cholesterol and an association between an allelic variant of the GPI-1 gene and low HDL. This analysis is based on comparing a pair of pools of DNA from individuals who exhibit respectively hypo or hypernormal levels of a biochemical trait (*e.g.*, cholesterol or HDL) and individually examining SNPs for a difference in allelic frequency between the pools. An association is considered to be positive if a statistically significant value of at least 3.841 using a 1-degree-of-freedom chi-squared test of association, $p = 0.05$, is obtained. Standard multiple testing corrections are applied if more than one SNP is considered at a time, i.e., multiple SNPs are tested during the same study. Although not always required, it may be necessary to further examine the frequency of allelic variants in other populations, including those exhibiting normal levels of the given trait.

For a qualitative trait (*e.g.*, hypertension) association studies are based on determining the occurrence of certain alleles in a given population of diseased *vs.* healthy individuals.

Allelic variants of COX6B, GPI-1 and other genes found to associate with high cholesterol, low HDL and/or cardiovascular disease can represent useful markers for indicating a predisposition for developing a risk factor for cardiovascular disease. These allelic variants may not necessarily represent functional variants affecting the expression, stability, or activity of the encoded protein product. Those of skill in the art would be able to determine which allelic variants are to be used, alone or in conjunction with other variants, only for indicating a predisposition for cardiovascular disease or for profiling of drug reactivity and for

determining those that may be also useful for screening for potential therapeutics.

Any method used to determine association can be used to discover or confirm the association of other polymorphic regions in the COX6B gene, the GPI-1 gene or any other gene that may be associated with cardiovascular disease.

G. Detection of Polymorphisms

1. Nucleic acid detection methods

Generally, these methods are based in sequence-specific polynucleotides, oligonucleotides, probes and primers. Any method known to those of skill in the art for detecting a specific nucleotide within a nucleic acid sequence or for determining the identity of a specific nucleotide in a nucleic acid sequence is applicable to the methods of determining the presence or absence of an allelic variant of a COX6B gene or GPI-1 gene or another gene associated with cardiovascular disease. Such methods include, but are not limited to, techniques utilizing nucleic acid hybridization of sequence-specific probes, nucleic acid sequencing, selective amplification, analysis of restriction enzyme digests of the nucleic acid, cleavage of mismatched heteroduplexes of nucleic acid and probe, alterations of electrophoretic mobility, primer specific extension, oligonucleotide ligation assay and single-stranded conformation polymorphism analysis. In particular, primer extension reactions that specifically terminate by incorporating a dideoxynucleotide are useful for detection. Several such general nucleic acid detection assays are described in U.S. Patent No. 6,030,778.

a. Primer extension-based methods

Several primer extension-based methods for determining the identity of a particular nucleotide in a nucleic acid sequence have been reported (see, *e.g.*, PCT Application No. PCT/US96/03651 (WO96/29431), PCT Application No. PCT/US97/20444 (WO 98/20019), PCT Application No. PCT/US91/00046 (WO91/13075), and U.S. Patent No. 5,856,092). In general, a primer is prepared that specifically hybridizes adjacent to a polymorphic site in a particular nucleic acid sequence. The primer is then extended in the presence of one or more dideoxynucleotides, typically with at least one of the dideoxynucleotides being the complement of the nucleotide that is polymorphic at the site. The primer and/or the dideoxynucleotides may be labeled to facilitate a determination of primer extension and identity of the extended nucleotide.

In one method, primer extension and/or the identity of the extended nucleotide(s) are determined by mass spectrometry (see, *e.g.*, PCT Application Nos. PCT/US96/03651 (WO96/29431) and PCT/US97/20444 (WO 98/20019)).

b. Polymorphism-specific probe hybridization

One exemplary detection method is allele specific hybridization using probes overlapping the polymorphic site and having about 5, 10, 15, 20, 25, or 30 nucleotides around the polymorphic region. The probes can contain aturally occurring or modified nucleotides (see U.S. Patent No. 6,156,501). For example, oligonucleotide probes may be prepared in which the known polymorphic nucleotide is placed centrally (allele-specific probes) and then hybridized to target DNA under conditions that permit hybridization only if a perfect match is found (Saiki *et al.* (1986) Nature 324:163; Saiki *et al.* (1989) Proc. Natl Acad. Sci USA 86:6230; and Wallace *et al.* (1979) Nucl. Acids Res. 6:3543). Such allele specific oligonucleotide hybridization techniques may be used for the simultaneous

detection of several nucleotide changes in different polymorphic regions. For example, oligonucleotides having nucleotide sequences of specific allelic variants are attached to a hybridizing membrane and this membrane is then hybridized with labeled sample nucleic acid. Analysis of the hybridization signal will then reveal the identity of the nucleotides of the sample nucleic acid. In one embodiment, several probes capable of hybridizing specifically to allelic variants are attached to a solid phase support, *e.g.*, a "chip". Oligonucleotides can be bound to a solid support by a variety of processes, including lithography. For example a chip can hold up to 250,000 oligonucleotides (GeneChip, Affymetrix, Santa Clara, CA). Mutation detection analysis using these chips comprising oligonucleotides, also termed "DNA probe arrays" is described *e.g.*, in Cronin *et al.* (1996) Human Mutation 7:244 and in Kozal *et al.* (1996) Nature Medicine 2:753. In one embodiment, a chip includes all the allelic variants of at least one polymorphic region of a gene. The solid phase support is then contacted with a test nucleic acid and hybridization to the specific probes is detected. Accordingly, the identity of numerous allelic variants of one or more genes can be identified in a simple hybridization experiment.

c. Nucleic acid amplification-based methods

In other detection methods, it is necessary to first amplify at least a portion of a COX6B gene, GPI-1 gene or another gene associated with cardiovascular disease prior to identifying the allelic variant. Amplification can be performed, *e.g.*, by PCR and/or LCR, according to methods known in the art. In one embodiment, genomic DNA of a cell is exposed to two PCR primers and amplification is performed for a number of cycles sufficient to produce the required amount of amplified DNA. In certain embodiments, the primers are located between 150 and 350 base pairs apart.

Alternative amplification methods include: self sustained sequence replication (Guatelli, J. C. *et al.* (1990) Proc. Natl. Acad. Sci. U.S.A. 87:1874-1878); transcriptional amplification system (Kwoh, D. Y. *et al.* (1989) Proc. Natl. Acad. Sci. U.S.A. 86:1173-1177); Q-Beta Replicase (Lizardi, P. M. *et al.* (1988) Bio/Technology 6:1197) and any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are also useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

Alternatively, allele specific amplification technology, which depends on selective PCR amplification may be used in conjunction with the alleles provided herein. Oligonucleotides used as primers for specific amplification may carry the allelic variant of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) Nucleic Acids Res. 17:2437-2448) or at the extreme 3′ end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) Tibtech 11:238; Newton *et al.* (1989) Nucl. Acids Res. 17:2503). In addition it may be desirable to introduce a restriction site in the region of the mutation to create cleavage-based detection (Gasparini *et al.* (1992) Mol. Cell Probes 6:1).

d. Nucleic acid sequencing-based methods

In one embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence at least a portion of the COX6B gene, GPI-1 gene or other gene associated with cardiovascular disease and to detect allelic variants, *e.g.*, mutations, by comparing the sequence of the sample sequence with the corresponding wild-type (control) sequence. Exemplary sequencing reactions include those based on techniques developed by Maxam and Gilbert (Proc. Natl. Acad. Sci.

USA (1977) 74:560) or Sanger (Sanger *et al.* (1977) Proc. Natl. Acad. Sci 74:5463). It is also contemplated that any of a variety of automated sequencing procedures may be used when performing the subject assays (Biotechniques (1995) 19:448), including sequencing by mass spectrometry (see, for example, U.S. Patent No. 5,547,835 and International PCT Application No. WO 94/16101, entitled DNA Sequencing by Mass Spectrometry by H. Koster; U.S. Patent No. 5,547,835 and International PCT Application No. WO 94/21822, entitled "DNA Sequencing by Mass Spectrometry Via Exonuclease Degradation" by H. Koster), and U.S. Pat. No. 5,605,798 and International Patent Application No. PCT/US96/03651 entitled DNA Diagnostics Based on Mass Spectrometry by H. Koster; Cohen *et al.* (1996) Adv Chromatogr 36:127-162; and Griffin *et al.* (1993) Appl Biochem Biotechnol 38:147-159). It will be evident to one skilled in the art that, for certain embodiments, the occurrence of only one, two or three of the nucleic acid bases need be determined in the sequencing reaction. For instance, A-track sequencing or an equivalent, *e.g.*, where only one nucleotide is detected, can be carried out. Other sequencing methods are disclosed, *e.g.*, in U.S. Patent No. 5,580,732 entitled "Method of DNA sequencing employing a mixed DNA-polymer chain probe" and U.S. Patent No. 5,571,676 entitled "Method for mismatch-directed *in vitro* DNA sequencing".

e. Restriction enzyme digest analysis

In some cases, the presence of a specific allele in nucleic acid, particularly DNA, from a subject can be shown by restriction enzyme analysis. For example, a specific nucleotide polymorphism can result in a nucleotide sequence containing a restriction site that is absent from the nucleotide sequence of another allelic variant.

f. Mismatch Cleavage

Protection from cleavage agents, such as, but not limited to, a nuclease, hydroxylamine or osmium tetroxide and with piperidine, can be used to detect mismatched bases in RNA/RNA DNA/DNA, or RNA/DNA heteroduplexes (Myers, *et al.* (1985) Science 230:1242). In general, the technique of "mismatch cleavage" starts by providing heteroduplexes formed by hybridizing a control nucleic acid, which is optionally labeled, *e.g.*, RNA or DNA, comprising a nucleotide sequence of an allelic variant with a sample nucleic acid, e.g, RNA or DNA, obtained from a tissue sample. The double-stranded duplexes are treated with an agent, which cleaves single-stranded regions of the duplex such as duplexes formed based on basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digest the mismatched regions.

In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine whether the control and sample nucleic acids have an identical nucleotide sequence or in which nucleotides they differ (see, for example, Cotton *et al.* (1988) Proc. Natl Acad Sci USA 85:4397; Saleeba *et al.* (1992) Methods Enzymod. 217:286-295). The control or sample nucleic acid is labeled for detection.

g. Electrophoretic mobility alterations

In other embodiments, alteration in electrophoretic mobility is used to identify the type of allelic variant in the COX6B gene, GPI-1 gene or other gene associated with cardiovascular disease. For example, single-strand conformation polymorphism (SSCP) may be used to detect

differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) Proc. Natl. Acad. Sci. USA 86:2766, see also Cotton (1993) Mutat Res 285:125-144; and Hayashi (1992) Genet Anal Tech Appl 9:73-79). Single-stranded DNA fragments of sample and control nucleic acids are denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In another embodiment, the subject method uses heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen *et al.* (1991) Trends Genet 7:5).

h. Polyacrylamide Gel Electrophoresis

In yet another embodiment, the identity of an allelic variant of a polymorphic region in the COX6B gene, GPI-1 gene or other gene associated with cardiovascular disease is obtained by analyzing the movement of a nucleic acid comprising the polymorphic region in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al.* (1985) Nature 313:495). When DGGE is used as the method of analysis, DNA will be modified to ensure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing agent gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) Biophys Chem 265:1275).

i. Oligonucleotide ligation assay (OLA)

In another embodiment, identification of the allelic variant is carried out using an oligonucleotide ligation assay (OLA), as described, *e.g.*, in U.S. Patent No. 4,998,617 and in Landegren, U. *et al.*, Science 241:1077-1080 (1988). The OLA protocol uses two oligonucleotides that are designed to be capable of hybridizing to abutting sequences of a single strand of a target. One of the oligonucleotides is linked to a separation marker, e.g,. biotinylated, and the other is detectably labeled. If the precise complementary sequence is found in a target molecule, the oligonucleotides will hybridize such that their termini abut, and create a ligation substrate. Ligation then permits the labeled oligonucleotide to be recovered using avidin, or another biotin ligand. Nickerson, D. A. *et al.* have described a nucleic acid detection assay that combines attributes of PCR and OLA (Nickerson, D. A. *et al.*, Proc. Natl. Acad. Sci. (U.S.A.) 87:8923-8927 (1990). In this method, PCR is used to achieve the exponential amplification of target DNA, which is then detected using OLA.

Several techniques based on this OLA method have been developed and can be used to detect specific allelic variants of a polymorphic region of a gene. For example, U.S. Pat. No. 5,593,826 discloses an OLA using an oligonucleotide having 3'-amino group and a 5'- phosphorylated oligonucleotide to form a conjugate having a phosphoramidate linkage. In another variation of OLA described in Tobe *et al.* (1996) Nucl. Acids Res. 24: 3728), OLA combined with PCR permits typing of two alleles in a single microtiter well. By marking each of the allele-specific primers with a unique hapten, i.e. digoxigenin and fluorescein, each OLA reaction can be detected by using hapten specific antibodies that are labeled with different enzyme reporters, alkaline phosphatase or horseradish peroxidase. This system permits the detection of the two alleles using a

high throughput format that leads to the production of two different colors.

j. SNP detection methods

Also provided are methods for detecting single nucleotide polymorphisms. Because single nucleotide polymorphisms constitute sites of variation flanked by regions of invariant sequence, their analysis requires no more than the determination of the identity of the single nucleotide present at the site of variation and it is unnecessary to determine a complete gene sequence for each patient. Several methods have been developed to facilitate the analysis of such single nucleotide polymorphisms.

In one embodiment, the single base polymorphism can be detected by using a specialized exonuclease-resistant nucleotide, as disclosed, *e.g.*, in Mundy, C. R. (U.S. Patent No. 4,656,127). According to the method, a primer complementary to the allelic sequence immediately 3′ to the polymorphic site is permitted to hybridize to a target molecule obtained from a particular animal or human. If the polymorphic site on the target molecule contains a nucleotide that is complementary to the particular exonuclease-resistant nucleotide derivative present, then that derivative will be incorporated onto the end of the hybridized primer. Such incorporation renders the primer resistant to exonuclease, and thereby permits its detection. Since the identity of the exonuclease-resistant derivative of the sample is known, a finding that the primer has become resistant to exonucleases reveals that the nucleotide present in the polymorphic site of the target molecule was complementary to that of the nucleotide derivative used in the reaction. This method has the advantage that it does not require the determination of large amounts of extraneous sequence data.

In another embodiment, a solution-based method for determining the identity of the nucleotide of a polymorphic site is employed (Cohen, D. *et al.* (French Patent 2,650,840; PCT Application No. WO91/02087)). As in the Mundy method of U.S. Patent No. 4,656,127, a primer is employed that is complementary to allelic sequences immediately 3′ to a polymorphic site. The method determines the identity of the nucleotide of that site using labeled dideoxynucleotide derivatives, which, if complementary to the nucleotide of the polymorphic site will become incorporated onto the terminus of the primer.

k. Genetic Bit Analysis

An alternative method, known as Genetic Bit Analysis or GBA™ is described by Goelet, *et al.* (U.S. Patent No. 6,004,744, PCT Application No. 92/15712). The method of Goelet, *et al.* uses mixtures of labeled terminators and a primer that is complementary to the sequence 3′ to a polymorphic site. The labeled terminator that is incorporated is thus determined by, and complementary to, the nucleotide present in the polymorphic site of the target molecule being evaluated. In contrast to the method of Cohen *et al.* (French Patent 2,650,840; PCT Application No. WO91/02087), the method of Goelet, *et al.* is typically a heterogeneous phase assay, in which the primer or the target molecule is immobilized to a solid phase.

l. Other primer-guided nucleotide incorporation procedures

Other primer-guided nucleotide incorporation procedures for assaying polymorphic sites in DNA have been described (Komher, J. S. *et al.*, Nucl. Acids Res. 17:7779-7784 (1989); Sokolov, B. P., Nucl. Acids Res. 18:3671 (1990); Syvanen, A. C., *et al.*, Genomics 8:684-692 (1990), Kuppuswamy, M. N. *et al.*, Proc. Natl. Acad. Sci. (U.S.A.) 88:1143-1147 (1991); Prezant, T. R. *et al.*, Hum. Mutat. 1:159-164

(1992); Ugozzoli, L. *et al.*, GATA 9:107-112 (1992); Nyren, P. *et al.*, Anal. Biochem. 208:171-175 (1993)). These methods differ from GBA™ in that they all rely on the incorporation of labeled deoxynucleotides to discriminate between bases at a polymorphic site. In such a format, since the signal is proportional to the number of deoxynucleotides incorporated, polymorphisms that occur in runs of the same nucleotide can result in signals that are proportional to the length of the run (Syvanen, A. C., *et al.*, Amer. J. Hum. Genet. 52:46-59 (1993)).

For determining the identity of the allelic variant of a polymorphic region located in the coding region of a gene, yet other methods than those described above can be used. For example, identification of an allelic variant that encodes a mutated protein can be performed by using an antibody specifically recognizing the mutant protein in, *e.g.*, immunohistochemistry or immunoprecipitation. Binding assays are known in the art and involve, *e.g.*, obtaining cells from a subject, and performing binding experiments with a labeled lipid, to determine whether binding to the mutated form of the protein differs from binding to the wild-type protein.

m. Molecular structure determination

If a polymorphic region is located in an exon, either in a coding or non-coding region of the gene, the identity of the allelic variant can be determined by determining the molecular structure of the mRNA, pre-mRNA, or cDNA. The molecular structure can be determined using any of the above described methods for determining the molecular structure of the genomic DNA, *e.g.*, sequencing and SSCP.

n. Mass spectrometric methods

Nucleic acids also can be analyzed by detection methods and protocols, particularly those that rely on mass spectrometry (see, *e.g.*, U.S. Patent No. 5,605,798, allowed co-pending U.S. Application Serial No. 08/617,256, allowed co-pending U.S. Application Serial No. 08/744,481, U.S. Application Serial No. 08/990,851, International PCT Application No. WO 98/20019). These methods can be automated (see, *e.g.*, co-pending U.S. Application Serial No. 09/285,481, which describes an automated process line). Among the methods of analysis herein are those involving the primer oligo base extension (PROBE) reaction with mass spectrometry for detection (described herein and elsewhere, see *e.g.*, U.S. Application Serial Nos. 08/617,256, 09/287,681, 09/287,682, 09/287,141 and 09/287,679, allowed co-pending U.S. Application Serial No. 08/744,481, International PCT Application No. PCT/US97/20444, published as International PCT Application No. WO 98/20019, and based upon U.S. Application Serial Nos. 08/744,481, 08/744,590, 08/746,036, 08/746,055, 08/786,988, 08/787,639, 08/933,792, 08/746,055, 08/786,988 and 08/787,639; see, also U.S. Application Serial No. 09/074,936, allowed U.S. Application Serial No. 08/787,639, and U.S. Application Serial Nos. 08/746,055 and 08/786,988, and published International PCT Application No. WO 98/20020).

One format for performing the analyses is a chip based format in which the biopolymer is linked to a solid support, such as a silicon or silicon-coated substrate, typicaly in the form of an addressable array. Typically when analyses are performed using mass spectrometry, particularly MALDI, nanoliter volumes of sample are loaded on, such that the resulting spot is about, or smaller than, the size of the laser spot. It has been found that when this is achieved, the results from the mass spectrometric analysis are quantitative. The area under the peaks in the

resulting mass spectra are proportional to concentration (when normalized and corrected for background). Methods for preparing and using such chips are described in allowed co-pending U.S. Application Serial No. 08/787,639, co-pending U.S. Application Serial Nos. 08/786,988, 09/364,774, 09/371,150 and 09/297,575; see, also U.S. Application Serial No. PCT/US97/20195, which published as International PCT Application No. WO 98/20020. Chips and kits for performing these analyses are commercially available from SEQUENOM under the trademark MassARRAY™. MassARRAY™ relies on the fidelity of the enzymatic primer extension reactions combined with the miniaturized array and MALDI-TOF (Matrix-Assisted Laser Desorption Ionization-Time of Flight) mass spectrometry to deliver results rapidly. It accurately distinguishes single base changes in the size of DNA fragments relating to genetic variants without tags.

Multiplex methods allow for the simultaneous detection of more than one polymorphic region in a particular gene or polymorphic regions in several genes. This is the one exemplary method for carrying out haplotype analysis of allelic variants of the COX6B and/or GPI-1 genes separately, or along with allelic variants of one or more other genes associated with cardiovascular disease.

Multiplexing can be achieved by several different methodologies. For example, several mutations can be simultaneously detected on one target sequence by employing corresponding detector (probe) molecules (*e.g.*, oligonucleotides or oligonucleotide mimetics). The molecular weight differences between the detector oligonucleotides must be large enough so that simultaneous detection (multiplexing) is possible. This can be achieved either by the sequence itself (composition or length) or by the introduction of mass-modifying functionalities into the detector oligonucleotides (see below).

Mass modifying moieties can be attached, for instance, to either the 5'-end of the oligonucleotide, to the nucleobase (or bases), to the phosphate backbone, and to the 2'-position of the nucleoside (nucleosides) and/or to the terminal 3'-position. Examples of mass modifying moieties include, for example, a halogen, an azido, or of the type, XR, wherein X is a linking group and R is a mass-modifying functionality. The mass-modifying functionality can thus be used to introduce defined mass increments into the oligonucleotide molecule.

The mass-modifying functionality can be located at different positions within the nucleotide moiety (see, *e.g.*, U.S. Patent No. 5,547,835 and International PCT Application No. WO 94/21822). For example, the mass-modifying moiety, M, can be attached either to the nucleobase, (in case of the c^7 -deazanucleosides also to C-7), to the triphosphate group at the alpha phosphate or to the 2'-position of the sugar ring of the nucleoside triphosphate. Modifications introduced at the phosphodiester bond, such as with alpha-thio nucleoside triphosphates, have the advantage that these modifications do not interfere with accurate Watson-Crick base-pairing and additionally allow for the one-step post-synthetic site-specific modification of the complete nucleic acid molecule *e.g.*, via alkylation reactions (see, *e.g.*, Nakamaye *et al.* (1988) Nucl. Acids Res. 16:9947-59). Exemplary mass-modifying functionalities are boron-modified nucleic acids since they are better incorporated into nucleic acids by polymerases (see, *e.g.*, Porter *et al.* (1995) Biochemistry 34:11963-11969; Hasan *et al.* (1996) Nucleic Acids Res. 24:2150-2157; Li *et al.* (1995) Nucl. Acids Res. 23:4495-4501).

Furthermore, the mass-modifying functionality can be added so as to affect chain termination, such as by attaching it to the 3'-position of the sugar ring in the nucleoside triphosphate. For those skilled in the art, it is clear that many combinations can be used in the methods provided

herein. In the same way, those skilled in the art will recognize that chain-elongating nucleoside triphosphates also can be mass-modified in a similar fashion with numerous variations and combinations in functionality and attachment positions.

For example, without being bound to any particular theory, the mass-modification can be introduced for X in XR as well as using oligo-/polyethylene glycol derivatives for R. The mass-modifying increment (m) in this case is 44, i.e. five different mass-modified species can be generated by just changing m from 0 to 4 thus adding mass units of 45 (m = 0), 89 (m = 1), 133 (m = 2), 177 (m = 3) and 221 (m = 4) to the nucleic acid molecule (*e.g.*, detector oligonucleotide (D) or the nucleoside triphosphates, respectively). The oligo/polyethylene glycols also can be monoalkylated by a lower alkyl such as, but are not limited to, methyl, ethyl, propyl, isopropyl and t-butyl. Other chemistries can be used in the mass-modified compounds (see, *e.g.*, those described in Oligonucleotides and Analogues, A Practical Approach, F. Eckstein, editor, IRL Press, Oxford, 1991).

In yet another embodiment, various mass-modifying functionalities, R, other than oligo/polyethylene glycols, can be selected and attached via appropriate linking chemistries, X. A simple mass-modification can be achieved by substituting H for halogens, such as F, Cl, Br and/or I, or pseudohalogens such as CN, SCN, NCS, or by using different alkyl, aryl or aralkyl moieties such as methyl, ethyl, propyl, isopropyl, t-butyl, hexyl, phenyl, substituted phenyl, benzyl, or functional groups such as CH_2F, CHF_2, CF_3, $Si(CH_3)_3$, $Si(CH_3)_2(C_2H_5)$, $Si(CH_3)(C_2H_5)_2$, $Si(C_2H_5)_3$. Yet another mass-modification can be obtained by attaching homo- or heteropeptides through the nucleic acid molecule (*e.g.*, detector (D)) or nucleoside triphosphates). One example, useful in generating mass-modified species with a mass increment of 57, is the attachment of

oligoglycines (m) to nucleic acid molecules (r), *e.g.*, mass-modifications of 74 (r = 1, m = 0), 131 (r = 1, m = 1), 188 (r = 1, m = 2), 245 (r = 1, m = 3) are achieved. Simple oligoamides also can be used, *e.g.*, but not limited to, mass-modifications of 74 (r = 1, m = 0), 88 (r = 2, m = 0), 102 (r = 3, m = 0), 116(r = 4, m = 0), are obtainable. Variations in additions to those set forth herein will be apparent to the skilled artisan.

Different mass-modified detector oligonucleotides can be used to simultaneously detect all possible variants/mutants simultaneously. Alternatively, all four base permutations at the site of a mutation can be detected by designing and positioning a detector oligonucleotide, so that it serves as a primer for a DNA/RNA polymerase with varying combinations of elongating and terminating nucleoside triphosphates. For example, mass modifications also can be incorporated during the amplification process.

A different multiplex detection format is one in which differentiation is accomplished by employing different specific capture sequences that are position-specifically immobilized on a flat surface (*e.g.*, a 'chip array'). If different target sequences T1-Tn are present, their target capture sites TCS1-TCSn will specifically interact with complementary immobilized capture sequences C1-Cn. Detection is achieved by employing appropriately mass differentiated detector oligonucleotides D1-Dn, which are mass modifying functionalities M1-Mn.

o. Other methods

Additional methods of analyzing nucleic acids include amplification-based methods including polymerase chain reaction (PCR), ligase chain reaction (LCR), mini-PCR, rolling circle amplification, autocatalytic methods, such as those using QJ replicase, TAS, 3SR, and any other suitable method known to those of skill in the art.

Other methods for analysis and identification and detection of polymorphisms, include but are not limited to, allele specific probes, Southern analyses, and other such analyses.

2. Primers and probes

Primers refer to nucleic acids that are capable of specifically hybridizing to a nucleic acid sequence that is adjacent to a polymorphic region of interest or to a polymorphic region and are extended. A primer can be used alone in a detection method, or a primer can be used together with at least one other primer or probe in a detection method. Primers also can be used to amplify at least a portion of a nucleic acid. For amplifying at least a portion of a nucleic acid, a forward primer (*i.e.*, 5′ primer) and a reverse primer (*i.e.*, 3′ primer) typically will be used. Forward and reverse primers hybridize to complementary stands of a double stranded nucleic acid, such that upon extension from each primer, a double stranded nucleic acid is amplified.

Probes refer to nucleic acids that hybridize to the region of interest and that are not further extended. For example, a probe is a nucleic acid that hybridizes adjacent to or at a polymorphic region of a COX6B gene, a GPI-1 gene or another gene associated with cardiovascular disease and that by hybridization or absence of hybridization to the DNA of a subject will be indicative of the identity of the allelic variant of the polymorphic region of the gene. Exemplary probes have a number of nucleotides sufficient to allow specific hybridization to the target nucleotide sequence. Where the target nucleotide sequence is present in a large fragment of DNA, such as a genomic DNA fragment of several tens or hundreds of kilobases, the size of a probe may have to be longer to provide sufficiently specific hybridization, as compared to a probe that is used to detect a target sequence that is present in a shorter fragment of DNA. For example, in some diagnostic methods, a portion of a COX6B

gene, a GPI-1 gene or another gene associated with cardiovascular disease may first be amplified and thus isolated from the rest of the chromosomal DNA and then hybridized to a probe. In such a situation, a shorter probe will likely provide sufficient specificity of hybridization. For example, a probe having a nucleotide sequence of about 10 nucleotides may be sufficient.

Exemplary primers and probes hybridize adjacent to or at the polymorphic sites described in TABLES 1-3. In addition, primers include SEQ ID NOS.: 5, 10, 43, 48, 53, 58, 63, 68, 73, 78, 83, 88, 93, 98, 103, 108, 113, and 118.

Primers and probes (RNA, DNA (single-stranded or double-stranded), PNA and their analogs) described herein may be labeled with any detectable reporter or signal moiety including, but not limited to radioisotopes, enzymes, antigens, antibodies, spectrophotometric reagents, chemiluminescent reagents, fluorescent and any other light producing chemicals. Additionally, these probes may be modified without changing the substance of their purpose by terminal addition of nucleotides designed to incorporate restriction sites or other useful sequences, proteins, signal generating ligands such as acridinium esters, and/or paramagnetic particles.

These probes may also be modified by the addition of a capture moiety (including, but not limited to para-magnetic particles, biotin, fluorescein, dioxigenin, antigens, antibodies) or attached to the walls of microtiter trays to assist in the solid phase capture and purification of these probes and any DNA or RNA hybridized to these probes. Fluorescein may be used as a signal moiety as well as a capture moiety, the latter by interacting with an anti-fluorescein antibody.

Any probe or primer can be prepared according to methods well known in the art and described, *e.g.*, in Sambrook, J. Fritsch, E.F., and

Maniatis, T. (1989(Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. For example, discrete fragments of the DNA can be prepared and cloned using restriction enzymes. Alternatively, probes and primers can be prepared using the Polymerase Chain Reaction (PCR) using primers having an appropriate sequence.

Oligonucleotides may be synthesized by standard methods known in the art, *e.g.* by use of an automated DNA synthesizer (such as are commercially available from, numerous sources, such as Biosearch (Novato, CA); and Applied Biosystems (Foster City, CA)). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein *et al.* ((1988) Nucl. Acids Res. 16:3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin *et al.*, 1988, Proc. Natl. Acad. Sci. U.S.A. 85:7448-7451), and others.

H. Transgenic Animals

Methods for making transgenic animals using a variety of transgenes are known (see, *e.g.*, Wagner *et al.* (1981) Proc. Nat. Acad. Sc. U.S.A. 78:5016; Stewart *et al.* (1982) Science 217:1046; Constantini *et al.* (1981) Nature 294:92; Lacy *et al.* (1982) Cell 34:343; McKnight *et al.* (1983) Cell 34:335; Brinstar *et al.* (1983) Nature 306:332; Palmiter *et al.* (1982) Nature 300:611; Palmiter *et al.* (1982) Cell 29:701 and Palmiter *et al.* (1983) Science 222:809; and U.S. Patent Nos. 6,175,057; 6,180,849 and 6,133,502).

Transgenic animals contain an exogenous nucleic acid sequence present as an extrachromosomal element or stably integrated in all or a portion of its cells, especially germ cells. Unless otherwise indicated, it will be assumed that a transgenic animal contains stable changes to the germline sequence. During the initial construction of the animal,

"chimeras" or "chimeric animals" are generated, in which only a subset of cells have the altered genome. Chimeras are primarily used for breeding purposes in order to generate the desired transgenic animal. Animals having a heterozygous alteration are generated by breeding of chimeras. Male and female heterozygotes are typically bred to generate homozygous animals.

The exogenous gene is usually either from a different species than the animal host, or is otherwise altered in its coding or non-coding sequence. The introduced gene may be a wild-type gene, naturally occurring polymorphism (*e.g.*, as described for COX6B, GPI-1 and other genes associated with cardiovascular disease) or a genetically manipulated sequence, for example having deletions, substitutions or insertions in the coding or non-coding regions. When the introduced gene is a coding sequence, it is usually operably linked to a promoter, which may be constitutive or inducible, and other regulatory sequences required for expression in the host animal.

Transgenic animals can contain other genetic alterations in addition to the presence of alleles of COX6B and/or GPI-1 genes. For example, the genome can be altered to affect the function of the endogenous genes, contain marker genes, or contain other genetic alterations (*e.g.*, alleles of other genes associated with cardiovascular disease).

A "knock-out" of a gene means an alteration in the sequence of the gene that results in a decrease of function of the target gene, typically such that target gene expression is undetectable or insignificant. A knock-out of an endogenous COX6B or GPI-1 gene means that function of the gene has been substantially decreased so that expression is not detectable or only present at insignificant levels. "Knock-out" transgenics can be transgenic animals having a heterozygous knock-out of the COX6B or GPI-1 gene or a homozygous knock-out of one or both of these genes.

"Knock-outs" also include conditional knock-outs, where alteration of the target gene can occur upon, for example, exposure of the animal to a substance that promotes target gene alteration, introduction of an enzyme that promotes recombination at the target gene site (*e.g.*, Cre in the Cre-lox system), or other method for directing the target gene alteration postnatally.

A "knock-in" of a target gene means an alteration in a host cell genome that results in altered expression (*e.g.*, increased (including ectopic)) of the target gene, *e.g.*, by introduction of an additional copy of the target gene, or by operatively inserting a regulatory sequence that provides for enhanced expression of an endogenous copy of the target gene. "Knock-in" transgenics of interest can be transgenic animals having a knock-in of the COX6B or GPI-1. Such transgenics can be heterozygous or homozygous for the knock-in gene. "Knock-ins" also encompass conditional knock-ins.

A construct is suitable for use in the generation of transgenic animals if it allows the desired level of expression of a COX6B or GPI-1 encoding sequence or the encoding sequence of another gene associated with cardiovascular disease. Methods of isolating and cloning a desired sequence, as well as suitable constructs for expression of a selected sequence in a host animal, are well known in the art and are described below.

For the introduction of a gene into the subject animal, it is generally advantageous to use the gene as a gene construct wherein the gene is ligated downstream of a promoter capable of and operably linked to expressing the gene in the subject animal cells. Specifically, a transgenic non-human mammal showing high expression of the desired gene can be created by microinjecting a vector ligated with said gene into a fertilized egg of the subject non-human mammal (*e.g.*, rat fertilized egg)

downstream of various promoters capable of expressing the protein and/or the corresponding protein derived from various mammals (rabbits, dogs, cats, guinea pigs, hamsters, rats, mice and other mammals)

Useful vectors include Escherichia coli-derived plasmids, Bacillus subtilis-derived plasmids, yeast-derived plasmids, bacteriophages such as lambda, phage, retroviruses such as Moloney leukemia virus, and animal viruses such as vaccinia virus or baculovirus.

Useful promoters for such gene expression regulation include, for example, promoters for genes derived from viruses (cytomegalovirus, Moloney leukemia virus, JC virus, breast cancer virus and others), and promoters for genes derived from various mammals (humans, rabbits, dogs, cats, guinea pigs, hamsters, rats, mice and other such mammalian species) and birds, such as, but are not limited to, chickens (*e.g.*, genes for albumin, insulin II, erythropoietin, endothelin, osteocalcin, muscular creatine kinase, platelet-derived growth factor beta, keratins K1, K10 and K14, collagen types I and II, atrial natriuretic factor, dopamine beta-hydroxylase, endothelial receptor tyrosine kinase (generally abbreviated Tie2), sodium-potassium adenosine triphosphorylase (generally abbreviated Na,K-ATPase), neurofilament light chain, metallothioneins I and IIA, metalloproteinase I tissue inhibitor, MHC class I antigen (generally abbreviated H-2L), smooth muscle alpha actin, polypeptide chain elongation factor 1 alpha (EF-1 alpha), beta actin, alpha and beta myosin heavy chains, myosin light chains 1 and 2, myelin base protein, serum amyloid component, myoglobin, renin and other such proteins.

The above-mentioned vectors can include a sequence for terminating the transcription of the desired messenger RNA in the transgenic animal (generally referred to as terminator); for example, gene expression can be manipulated using a sequence with such function contained in various genes derived from viruses, mammals and birds. The

simian virus SV40 terminator is a commonly used exemplary terminator. Additionally, for the purpose of increasing the expression of the desired gene, the splicing signal and enhancer region of each gene, a portion of the intron of a eukaryotic organism gene may be ligated 5′ upstream of the promoter region, or between the promoter region and the translational region, or 3′ downstream of the translational region as desired.

A translational region for a protein of interest can be obtained using the entire or portion of genomic DNA of blood, kidney or fibroblast origin from various mammals (humans, rabbits, dogs, cats, guinea pigs, hamsters, rats, mice and others) or of various commercially available genomic DNA libraries, as a starting material, or using complementary DNA prepared by a known method from RNA of blood, kidney or fibroblast origin as a starting material. Also, an exogenous gene can be obtained using complementary DNA prepared by a known method from RNA of human fibroblast origin as a starting material. All these translational regions can be used in transgenic animals.

To obtain the translational region, it is possible to prepare DNA incorporating an exogenous gene encoding the protein of interest in which the gene is ligated downstream of the above-mentioned promoter (generally upstream of the translation termination site) as a gene construct capable of being expressed in the transgenic animal.

DNA constructs for random integration need not include regions of homology to mediate recombination. Where homologous recombination is desired, the DNA constructs contain at least a portion of the target gene with the desired genetic modification, and include regions of homology to the target locus. Conveniently, markers for positive and negative selection are included. Methods for generating cells having targeted gene modifications through homologous recombination are known in the art.

For various techniques for transfecting mammalian cells, see Keown *et al.* (1990) Methods in Enzymology 185:527-537.

The transgenic animal can be created by introducing a COX6B or GPI-1 gene construct into, for example, an unfertilized egg, a fertilized egg, a spermatozoon or a germinal cell containing a primordial germinal cell thereof, generally in the embryogenic stage in the development of a non-human mammal (typically in the single-cell or fertilized cell stage and generally before the 8-cell phase), by standard means, such as the calcium phosphate method, the electric pulse method, the lipofection method, the agglutination method, the microinjection method, the particle gun method, the DEAE-dextran method and other such method. Also, it is possible to introduce a desired COX6B or GPI-1 gene into a, for example, somatic cell, a living organ, a tissue cell, for example, by gene transformation methods, and use it for cell culture, tissue culture and other such uses. Furthermore, these cells may be fused with the above-described germinal cell by a commonly known cell fusion method to create a transgenic animal.

For embryonic stem (ES) cells, an ES cell line may be employed, or embryonic cells may be obtained freshly from a host, *e.g.* mouse, rat, guinea pig, and other mammals and birds. Such cells are grown on an appropriate fibroblast-feeder layer or grown in the presence of appropriate growth factors, such as leukemia inhibiting factor (LIF). When ES cells have been transformed, they may be used to produce transgenic animals. After transformation, the cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be detected by employing a selective medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of homologous recombination or integration of the construct. Those colonies that are positive may then be used for embryo manipulation and blastocyst

injection. Blastocysts are obtained from 4 to 6 week old superovulated females. The ES cells are trypsinized, and the modified cells are injected into the blastocoel of the blastocyst. After injection, the blastocysts are returned to each uterine horn of pseudopregnant females. Females are then allowed to go to term and the resulting litters screened for mutant cells having the construct. By providing for a different phenotype of the blastocyst and the ES cells, chimeric progeny can be readily detected. The chimeric animals are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs can be maintained as allogeneic or congenic grafts or transplants, or in *in vitro* culture.

Animals containing more than one transgene, such as allelic variants of COX6B and/or GPI-1 and/or other genes associated with cardiovascular disease can be made by sequentially introducing individual alleles into an animal in order to produce the desired phenotype (manifestation or predisposition to cardiovascular disease).

I. Effect of Allelic Variants on the Encoded Protein and Disease Related Phenotype

The effect of an allelic variant on a COX6B or GPI-1 protein (altered amount, stability, location and/or activity) can be determined according to methods known in the art. Allelic variants of the COX6B and GPI-1 genes can be assayed individually or in combination with other variants known to be associated with cardiovascular disease.

If the mutation is located in an intron, the effect of the mutation can be determined, *e.g.*, by producing transgenic animals in which the allelic variant linked to lipid metabolism and/or cardiovascular disease has been introduced and in which the wild-type gene or predominant allele may have been knocked out. Comparison of the level of expression of the

protein in the mice transgenic for the allelic variant with mice transgenic for the predominant allele will reveal whether the mutation results in increased or decreased synthesis of the associated protein and/or aberrant tissue distribution of the associated protein. Such analysis could also be performed in cultured cells, in which the human variant allele gene is introduced and, *e.g.*, replaces the endogenous gene in the cell. Thus, depending on the effect of the alteration a specific treatment can be administered to a subject having such a mutation. Accordingly, if the mutation results in decreased production of a COX6B or GPI-1 protein, the subject can be treated by administration of a compound that increases synthesis, such as by increasing COX6B or GPI-1 gene expression, and wherein the compound acts at a regulatory element different from the one that is mutated. Alternatively, if the mutation results in increased COX6B or GPI-1 protein levels, the subject can be treated by administration of a compound that reduces protein production, *e.g.*, by reducing COX6B or GPI-1 gene expression or a compound that inhibits or reduces the activity of COX6B or GPI-1 protein.

J. Diagnostic and Prognostic Assays

Typically, an individual allelic variant that associates with a risk factor for cardiovascular disease will not be used in isolation as a prognosticator for a subject developing high cholesterol, low HDL or cardiovascular disease. An allelic variant typically will be one of a plurality of indicators that are used. The other indicators may be the manifestation of other risk factors for cardiovascular disease, *e.g.*, family history, high blood pressure, weight, activity level and other indicators, or additional allelic variants in the same or other genes associated with altered lipid metabolism and/or cardiovascular disease.

Useful combinations of allelic variants of the COX6B gene and/or the GPI-1 gene can be determined by examining combinations of variants

of these genes, which are assayed individually or assayed simultaneously using multiplexing methods as described above or any other labelling method that allows different variants to be identified. In particular, variants of COX6B gene and/or the GPI-1 gene may be assayed using kits (see below) or any of a variety microarrays known to those in the art. For example, oligonucleotide probes comprising the polymorphic regions surrounding any polymorphism in the COX6B or GPI-1 gene may be designed and fabricated using methods such as those described in U.S. Patent Nos. 5,492,806; 5,525,464; 5,695,940; 6,018,041; 6,025,136; WO 98/30883; WO 98/56954; WO99/09218; WO 00/58516; WO 00/58519, or references cited therein. Similarly one of skill in the art can determine useful combinations of allelic variants of the COX6B and/or GPI-1 genes along with variants of other genes associated with cardiovascular disease.

K. Pharmacogenomics

Subjects having one or more different allelic variants of the COX6B or GPI-1 polymorphic regions will respond differently to therapeutic drugs to treat cardiovascular disease or conditions. For example, there are numerous drugs available for lowering cholesterol levels: including lovastatin (MEVACOR; Merck & Co.), simvastatin (XOCOR; Merck & Co.), dextrothyroxine (CHOLOXIN; Knoll Pharmaceutical Co.), pamaqueside (Pfizer), cholestryramine (QUESTRAN; Bristol-Myers Squibb), colestipol (COLESTID; Pharmacia & Upjohn), acipomox (Pharmacia & Upjohn), fenofibrate (LIPIDIL), gemfibrozil (LOPID; Warner-Lambert), cerivastatin (LIPOBAY; Bayer), fluvastatin (LESCOL; Novartis), atorvastatin (LIPITOR, Warner-Lambert), etofylline clofibrate (DUOLIP; Merckle (Germany)), probucol (LORELCO; Hoechst Marion Roussel), omacor (Pronova (Norway), etofibrate (Merz (Germany), clofibrate (ATROMID-S; Wyeth-Ayerst (AHP)), and niacin (numerous manufacturers). All patients do not

respond identically to these drugs. Alleles of the COX6B or the GPI-1 gene that associate with altered lipid metabolism will be useful alone or in conjunction with markers in other genes associated with the development of cardiovascular disease to predict a subject's response to a therapeutic drug. For example, multiplex primer extension assays or microarrays comprising probes for alleles are useful formats for determining drug response. A correlation between drug responses and specific alleles or combinations of alleles of the COX6B or GPI-1 genes and other genes associated with cardiovascular disease can be shown, for example, by clinical studies wherein the response to specific drugs of subjects having different allelic variants of polymorphic regions of the COX6B or GPI-1 genes alone or in combination with allelic variants of other genes are compared. Such studies also can be performed using animal models, such as mice having various alleles and in which, *e.g.*, the endogenous COX6B or GPI-1 genes have been inactivated such as by a knock-out mutation. Test drugs are then administered to the mice having different alleles and the response of the different mice to a specific compound is compared. Accordingly, assays, microarrays and kits are provided for determining the drug that will be best suited for treating a specific disease or condition in a subject based on the individual's genotype. For example, it will be possible to select drugs that will be devoid of toxicity, or have the lowest level of toxicity possible for treating a subject having a disease or condition, *e.g.*, cardiovascular disease or high cholesterol or low HDL.

L. Kits

Kits can be used to indicate whether a subject is at risk of developing high cholesterol, low HDL and/or cardiovascular disease. The kits also can be used to determine if a subject who has high cholesterol or low HDL carries associated variants in the COX6B or GPI-1 genes or other

cardiovascular disease-related genes. This information could be used, *e.g.*, to optimize treatment of such individuals as a particular genotype may be associated with drug response.

In certain, the kits include a probe or primer that is capable of hybridizing adjacent to or at a polymorphic region of a COX6B or GPI-1 gene and thereby identifying whether the COX6B or GPI-1 gene contains an allelic variant that is associated with cardiovascular disease. Primers or probes that specifically hybridize at or adjacent to the SNPs described in Tables 1-3 could be included. In particular, primers or probes that contain the sequences of SEQ ID NOs.: 5, 10, 43, 48, 53, 58, 63, 68, 73, 78, 83, 88, 93, 98, 103, 108, 113, and 118 could be included in the kits. The kits optionally also include instructions for use in carrying out assays, interpreting results and diagnosing a subject as having a predisposition toward developing high cholesterol, low HDL and/or cardiovascular disease.

Exemplary kits for amplifying a region of a COX6B gene, GPI-1 gene, or other genes associated with cardiovascular disease (such as those listed in Table 3) contain two primers that flank a polymorphic region of the gene of interest. For example primers can include the sequences of SEQ ID NOs.: 3, 4, 8, 9, 41, 42, 46, 47, 51, 52, 56, 57, 61, 62, 66, 67, 71, 72, 76, 77, 81, 82, 86, 87, 91, 92, 96, 97, 101, 102, 106, 107, 111, 112, 116, and 117. For other assays, primers or probes hybridize to a polymorphic region or 5′ or 3′ to a polymorphic region depending on which strand of the target nucleic acid is used. For example, specific probes and primers contain sequences designated as SEQ ID NOs: 5, 10, 43, 48, 53, 58, 63, 68, 73, 78, 83, 88, 93, 98, 103, 108, 113, and 118. Those of skill in the art can synthesize primers and probes that hybridize adjacent to or at the polymorphic regions

described in TABLES 1-3 and other SNPs in genes associated with cardiovascular disease.

Yet other kits contain at least one reagent necessary to perform an assay. For example, the kit can comprise an enzyme, such as a nucleic acid polymerase. Alternatively the kit can contain a buffer or any other necessary reagent.

Yet other kits contain microarrays of probes to detect allelic variants of COX6B, GPI-1, and other genes associated with cardiovascular disease. The kits further contain instructions for their use and interpreting the results.

The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention. The practice of methods and development of the products provided herein employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See, for example, Molecular Cloning A Laboratory Manual, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); DNA Cloning, Volumes I and II (D.N. Glover ed., 1985); Oligonucleotide Synthesis (M.J. Gait ed., 1984); Mullis *et al.* U.S. Patent No. 4,683,195; Nucleic Acid Hybridization (B.D. Hames & S.J. Higgins eds. 1984); Transcription and Translation (B.D. Hames & S.J. Higgins eds. 1984); Culture of Animal Cells (R.I. Freshney, Alan R. Liss, Inc., 1987); Immobilized Cells and Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide To Molecular Cloning (1984); the treatise, Methods In Enzymology (Academic Press, Inc., New York); Gene Transfer Vectors For Mammalian Cells (J.H. Miller and M.P. Calos eds., 1987, Cold Spring Harbor Laboratory); Methods In Enzymology, Vols. 154 and 155 (Wu *et al.* eds., Immunochemical

Methods In Cell and Molecular Biology (Mayer and Walker, eds., Academic Press, London, 1987); Handbook of Experimental Immunology, Volumes I-IV (D.M. Weir and C.C. Blackwell, eds., 1986); Manipulating the Mouse Embryo, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986).

EXAMPLE 1

Isolation of DNA from blood samples of a stratified population

Blood samples were obtained from a population of unrelated Caucasian women between the ages of 18-79 (average age = 48). The women had, no response to media campaigns, attended the Twin Research Unit at the St. Thomas Hospital in London, England. For current purposes, only one member of a twin pair was used to insure that all observations were independent. Blood samples from 1400 unrelated individuals were measured for levels of cholesterol and HDL. Cholesterol and HDL level in blood samples were quantitated using standard assay methods.

The population was stratified into pools of 200 people, which represented the lower extreme and the upper extreme for serum levels of cholesterol and HDL.

<u>Cholesterol</u>

Pool 1: Individuals were considered to have low cholesterol (0.12 - 3.6 mmoles/L).

Pool 2: Individuals were considered to have high cholesterol (5.25 - 11.57 mmoles/L).

HDL

Pool 3:	Individuals were considered to have low levels of HDL (0.240 - 1.11 mmoles/L)
Pool 4:	Individuals were considered to have high levels of HDL (2.10 - 3.76 mmoles/L).

DNA extraction protocol

DNA was extracted from blood samples of each of the pools by utilizing the following protocol.

Section 1

1. Blood was extracted into EDTA tubes.
2. Blood sample was spun at 3,000 rpm for 10 minutes in a clinical centrifuge.
3. The buffy coat (the leucocytes, a yellowish layer of cells on top of the red blood cells) was removed and pooled into a 1 ml conical tube.
4. 0.9% saline was added to fill the tube and resuspend the leucocytes. Sample were immediately further processed or stored at 4°C for 24 hrs.
5. The sample was spun at 2,500 rpm for 10 minutes.
6. The buffy coat was again removed as cleanly as possible leaving behind any red cells, the sample was suspended in red cell lysis buffer and left for 20 minutes at 4°C.
7. The sample was spun again at 2,500 rpm for 10 minutes. If a pellet of unlysed red cells remained lying above the leucocytes the treatment with red cell lysis buffer was repeated.
8. The leucocyte pellet was resuspended in 2 ml 0.9% saline.
9. The DNA was liberated by the addition of leucocyte lysis buffer - the tube was capped and gently inverted several

times, until the liquid became viscous with DNA. The samples were handled with care to avoid shearing and damage to the DNA.

10. Samples were frozen for storage prior to full extraction.

Section 2

11. 2 ml of 5 M sodium perchlorate was added to the thawed sample and mixed by inversion. The sample was heated to 60°C for 30 - 40 minutes to fully denature proteins.
12. An equal volume of chloroform/isoamyl alcohol (24:1) was added at room temperature and the sample mixed for 10 minutes.
13. The sample was spun without a break at 3,000 rpm for 10 minutes.
14. The top aqueous phase was removed into a clean tube and two volumes of cold 100% ethanol added and mixed by inversion to precipitate DNA.
15. The DNA was removed using a sterile loop and resuspended in 1-5 ml TE buffer depending on the DNA yield.
16. The optical density was measured at 260 and 280 nm to check yield and purity of the DNA sample. For use in Examples 2 and 3, all DNA had an absorbance ratio of 1.6 at 260/280, a total yield of 32 μg and a concentration of 10 ng/μl. If initial purity levels were unacceptable a re-extraction was carried out (sections 12-15 above).

EXAMPLE 2

Detection of an Association Between an SNP at Position 86 of the Human COX6B Gene and High Cholesterol

DNA samples (as prepared in Example 1), representing 200 women, from the lower extreme, pool 1 (low levels of cholesterol) and the upper extreme, pool 2 (high levels of cholesterol) were amplified and analyzed for genetic differences using a MassEXTEND™ assay detection method. For each pool, single nucleotide polymorphisms were examined throughout the entire genome to detect differences in allelic frequency of a variant allele between the pools.

PCR Amplification of Samples from Pools 1 and 2

PCR primers were synthesized by Operon (Alameda, CA) using phosphoramidite chemistry. Amplification of the COX6B target sequence was carried out in two 50 μl PCR reactions with 100 ng of pooled human genomic DNA, obtained as described in Example 1, taken from samples in pool 1 or pool 2, although amounts ranging from 100 ng to 1 ug could be used. Individual DNA concentrations within the pooled samples were present in equal concentration with a final concentration of 0.5 ng. Each reaction contained 1X PCR buffer (Qiagen, Valencia, CA), 200 μM dNTPs, 1U Hotstar Taq polymerase (Qiagen, Valencia, CA), 4 mM $MgCl_2$, and 25 pmols of the long primer containing both the universal primer sequence and the target specific sequence 5′-AGCGGATAACAATTTCACACAGGTAGTCTGGTTCTGGTTGGGG-3′ (SEQ ID NO.: 4) , 2 pmoles of the short primer 5′-AGGATTCAGCACCATGGC-3′ (SEQ ID NO.: 3) and I0 pmoles of a biotinylated universal primer complementary to the 5′ end of the PCR amplicon 5′-AGCGGATAACAATTTCACACAGG-3′ (SEQ ID NO.: 121). Alternatively, the biotinylated universal primer could be 5′-GGCGCACGCCTCCACG-3′ (SEQ ID NO.: 122). After an initial round of

amplification with the target with the specific forward (long) and reverse primer (short), the 5′ biotinylated universal primer then hybridized and acted as a reverse primer thereby introducing a 3′ biotin capture moiety into the molecule. The amplification protocol results in a 5′-biotinylated double stranded DNA amplicon and dramatically reduces the cost of high throughput genotyping by eliminating the need to 5′ biotin label each forward primer used in a genotyping. Thermal cycling was performed in 0.2 mL tubes or 96 well plate using an MJ Research Thermal Cycler (Waltham, MA) (calculated temperature) with the following cycling parameters: 94°C for 5 min; 45 cycles: 94°C for 20 sec, 56°C for 30 sec, 72°C for 60 sec; 72°C 3 min.

Immobilization of DNA

The 50μl PCR reaction was added to 25μl of streptavidin coated magnetic bead (Dynal, Lake Success, NY) prewashed three times and resuspended in 1 M NH_4Cl, 0.06 M NH_4OH. The PCR amplicons were allowed to bind to the beads for 15 minutes at room temperature. The beads were then collected with a magnet and the supernatant containing unbound DNA was removed. The unbound strand was released from the double stranded amplicons by incubation in 100 mM NaOH and washing of the beads three times with 10 mM Tris pH 8.0.

Genotyping

The frequency of the alleles at position 86 in the COX6B gene was measured using the MassEXTEND™ assay and MALDI-TOF. The SNP identified at position 86 of COX6B in the GenBank sequence is represented as a C to T transversion. The MassEXTEND™ assay used detected the sequence of the complementary strand, thus the SNP was represented as G to A in the primer extension products. The DNA coated magnetic beads were resuspended in 26 mM Tris-HCL pH 9.5, 6.5 mM $MgCl_2$ and 50 mM each of dTTPs and 50 mM each of ddCTP, ddATP,

ddGTP, 2.5 U of a thermostable DNA polymerase (Amersham Pharmacia Biotech, Piscataway, NJ) and 20 pmoles of a template specific oligonucleotide primer 5'-AATCAAGAACTACAAGAC-3' (SEQ ID NO.: 5) (Operon, Alameda, CA). Primer extension occurred with three cycles of oligonucleotide primer hybridization and extension. The extension products were analyzed after denaturation from the template with 50 mM NH_4Cl and transfer of 150 nl of each sample to a silicon chip preloaded with 150 nl of H3PA (3-hydroxy picolinic acid) (Sigma Aldrich, St Louis, MO) matrix material. The sample material was allowed to crystallize and analyzed by MALDI-TOF (Bruker Daltonics, Billerica, MA; PerSeptive, Foster City, CA). The mass of the primer used in the MassEXTEND™ reaction was 5493.70 daltons. The predominant allele is extended by the addition of ddC, which has a mass of 5766.90 daltons. The allelic variant results in the addition of dT and ddG to the primer to produce an extension product having a mass of 6111.10 daltons.

In addition to being analyzed as part of a pool, each individual sample (0.5 ng) was amplified as described above and analyzed individually using a MassEXTEND™ reaction as described above.

Pooled populations of women (200 women per pool) with high cholesterol (pool 2) showed an increase in the frequency of the A allele at nucleotide position 86 of COX6B as compared with those with low levels of cholesterol (pool 1) (see Fig. 1). The association of this allelic variant of the COX6B gene with high cholesterol gave a statistically significant value of 14.30 using a 1-degree-of-freedom chi-squared test of association. In other words, the increase of 2.75% to 9.05% is significant, with a p value of 0.000156 (see Fig. 1). The genotype of each of the individuals in the pooled population was also determined by carrying out MassEXTEND™ reactions on each DNA samples individually. These analysis confirmed the pooling data showing that there was an

increase in the frequency of the A allele of 2.27% to 9.93%, (p = 0.0000061). The genotypes in pool 2 showed a decrease in the homozygous GG genotype from 95.4% to 82.35% and an increase in the heterozygous GA genotype from 4.55% to 15.44%. None of the individuals with low levels of serum cholesterol exhibited the homozygous AA genotype.

EXAMPLE 3

Detection of an Association Between an SNP at Position 2577 of the Human GPI-1 Gene and Low HDL

DNA samples (as prepared in Example 1), representing 200 women, from pool 3 (low level of HDL) and pool 4 (high levels of HDL) were amplified and analyzed for genetic differences using a MassEXTEND™ detection method. For each pool, SNPs were examined throughout the genome to detect differences in allelic frequency of variant alleles between the pools.

<u>PCR Amplification of Samples from Pools 3 and 4</u>

PCR primers were synthesized by Operon (Alameda, CA) using phosphoramidite chemistry. Amplification of the GPI-1 target sequence was carried out in single 50μl PCR reaction with 100 ng of pooled human genomic DNA (200 samples), obtained as described in Example 1, taken from samples in pool 3 or pool 4, although amounts ranging from 100 ng to 1 ug could be used. Individual DNA concentrations within the pooled samples were present in equal concentration with the final concentration of 0.5 ng. Each reaction contained 1X PCR buffer (Qiagen, Valencia, CA), 200 uM dNTPs, 1U Hotstar Taq polymerase (Qiagen, Valencia, CA), 4 mM $MgCl_2$, and 25 pmols of the forward primer containing both the universal primer sequence and the target specific short sequence 5'-AGCAGGGCTTCCTCCTTC-3' (SEQ ID NO.: 8) 2 pmoles of the long

primer 5'-AGCGGATAACAATTTCACACAGGTGACCCAGCCGTACCTATTC-3' (SEQ ID NO.: 9) and 10 pmoles of a biotinylated universal primer complementary to the 5' end of the PCR amplicon 5'-AGCGGATAACAATTTCACACAGG-3' (SEQ ID NO.: 121). After an initial round of amplification with the target with the specific forward (long) and reverse primer (short), the 5' biotinylated universal primer then hybridized and acted as a reverse primer thereby introducing a 3' biotin capture moiety into the molecule. The amplification protocol results in a 5'-biotinylated double stranded DNA amplicon and dramatically reduces the cost of high throughput genotyping by eliminating the need to 5' biotin label each forward primer used in a genotyping. Thermal cycling was performed in 0.2 mL tubes or 96 well plate using an MJ Research Thermal Cycler (Watham, MA) (calculated temperature) with the following cycling parameters: 94°C for 5 min; 45 cycles: 94°C for 20 sec, 56°C for 30 sec, 72°C for 60 sec; 72°C 3 min.

<u>Immobilization of DNA</u>

The 50 μl PCR reaction was added to 25 μl of streptavidin coated magnetic bead (Dynal, Lake Success, NY) prewashed three times and resuspended in 1 M NH_4Cl, 0.06 M NH_4OH. The PCR amplicons were allowed to bind to the beads for 15 minutes at room temperature. The beads were then collected with a magnet and the supernatant containing unbound DNA was removed. The unbound strand was released from the double stranded amplicons by incubation in 100 mM NaOH and washing of the beads three times with 10 mM Tris pH 8.0.

<u>Genotyping</u>

The frequency of the alleles at position 2577 in the GPI-1 gene was measured using the MassEXTEND™ assay and MALDI-TOF. The SNP identified at position 2577 of GPI-1 in the GenBank sequence is represented as a G to A transversion. The MassEXTEND™ assay used

detected this sequence, thus the SNP was represented as C to T in the primer extension products. The DNA coated magnetic beads were resuspended in 26 mM Tris-HCL pH 9.5, 6.5 mM $MgCl_2$ and 50 mM each of dTTPs and 50 mM each of ddCTP, ddATP, ddGTP, 2.5 U of a thermostable DNA polymerase (Amersham Pharmacia Biotech, Piscataway, NJ) and 20 pmoles of a template specific oligonucleotide primer 5'-AAGGGAGACAGATTTGGC-3' (SEQ ID NO.: 10) (Operon, Alameda, CA). Primer extension occurred with three cycles of oligonucleotide primer hybridization and extension. The extension products were analyzed after denaturation from the template with 50 mM NH_4Cl and transfer of 150 nl each sample to a silicon chip preloaded with 150 nl of H3PA matrix material. The sample material was allowed to crystallize and analyzed by MALDI-TOF (Bruker Daltonics, Billerica, MA; PerSeptive, Foster City, CA). The mass of the primer used in the MassEXTEND™ reaction was 5612.70 daltons. The predominant allele is extended by the addition of ddC, which has a mass of 5885.90 daltons. The allelic variant results in the addition of dT and ddG to the primer to produce an extension product having a mass of 6230.10 daltons.

In addition to being analyzed as a pool, each individual sample (0.5 ng) was amplified as described above and analyzed individually using the MassEXTEND™ reaction as described above.

Pooled populations of women (200 women per pool) with low HDL (pool 3) showed an increase in the T allele of 11.33% at nucleotide position 2577 as compared with those with high levels of HDL (pool 4). The association of this allelic variant of the GPI-1 gene with low HDL gave a statistically significant value of 15.04 using a 1-degree-of-freedom chi-squared test of association. In other words, the increase of 16.23% to 27.57% is significant, with a p value of 0.0001064 (see Fig. 2). The genotype of each of the individuals in the pooled population was also

determined by carrying out individual MassEXTEND™ reactions on individual DNA samples. These analysis confirmed the pooling data showing that there was an increase in the frequency of the T allele of 19.49% to 26.1%, (p=0.024). The measured genotypes in pool 3 showed a decrease in the homozygous CC genotype from 65.24% to 54.21% and an increase in the heterozygous CT genotype from 30.51% to 39.25%. The homozygous TT genotypes increased 2.3%.

Since modifications will be apparent to those of skill in this art, it is intended that this invention be limited only by the scope of the appended claims.

WHAT IS CLAIMED:

1. A method for detecting the presence or absence in a subject of at least one allelic variant of a polymorphic region of a gene associated with cardiovascular disease, comprising:

the step of detecting the presence or absence of an allelic variant of a polymorphic region of a cytochrome C oxidase subunit VIb (COX6B) gene of the subject that is associated with high serum cholesterol or an allelic variant of a polymorphic region of a N-acetylglucosaminyl transferase component (GPI-1) gene of the subject that is associated with low serum high density lipoprotein (HDL).

2. The method of claim 1, wherein the allelic variant is of a polymorphic region of the cytochrome C oxidase subunit VIb (COX6B) gene.

3. The method of claim 1, wherein the allelic variant is of a polymorphic region of the N-acetylglucosaminyl transferase component (GPI-1) gene.

4. The method of any of claims 1-3, further comprising detecting the presence or absence in a subject of least one allelic variant of another gene associated with cardiovascular disease.

5. The method of claim 4, wherein the other gene is selected from the group consisting of cholesterol ester transfer protein, plasma (CETP); apolipoprotein A-IV (APO A4); apolipoprotein A-I (APO A1); apolipoprotein E (APO E); apolipoprotein B (APO B); apolipoprotein C-III (APO C3); a gene encoding lipoprotein lipase (LPL); ATP-binding cassette transporter (ABC 1); paraoxonase 1 (PON 1); paraoxonase 2 (PON 2); 5,10-methylenetetrahydrofolate r reductase (MTHFR); a gene encoding hepatic lipase, E-selectin, G protein beta 3 subunit and angiotensin II type 1 receptor gene.

6. The method of claim 2 or claim 3, wherein the polymorphic region is a single nucleotide polymorphism (SNP).

7. The method of any of claims 1-6, wherein the detection is effected by detecting a a light producing reagent.

8. The method of claim 6, wherein the SNP is at position 86 of the cytochrome C oxidase subunit VIb (COX6B) gene coding sequence and the allelic variant is represented by a T nucleotide in the sense strand or an A nucleotide in the corresponding position in the antisense strand.

9. The method of claim 6, wherein the SNP is at position 2577 of the N-acetylgluocsaminyl transferase component GPI-1 (GPI-1) gene sequence and the allelic variant is represented by an A nucleotide in the sense strand or a T nucleotide in the corresponding position in the antisense strand.

10. The method of any of claims 1-3, wherein the detecting step is by a method selected from the group consisting of allele specific hybridization, primer specific extension, oligonucleotide ligation assay, restriction enzyme site analysis and single-stranded conformation polymorphism analysis.

11. The method of claim 8, further comprising:

(a) hybridizing a target nucleic acid comprising a cytochrome C oxidase subunit VIb (COX6B)-encoding nucleic acid or fragment thereof with a nucleic acid primer that hybridizes adjacent to nucleotide 86 of the coding sequence of the COX6B gene;

(b) extending the nucleic acid primer using the target nucleic acid as a template; and

(c) determining the mass of the extended primer to identify the nucleotide present at position 86, thereby determining the presence or absence of the allelic variant.

12. The method of claim 9, further comprising:

(a) hybridizing a target nucleic acid comprising a N-acetylglucosaminyl transferase component GPI-1 (GPI-1)-encoding

nucleic acid or fragment thereof with a nucleic acid primer that hybridizes adjacent to nucleotide 2577 of the GPI-1 gene;

(b) extending the nucleic acid primer using the target nucleic acid as a template; and

(c) determining the mass of the extended primer to identify the nucleotide present at position 2577, thereby determining the presence or absence of the allelic variant.

13. The method of any of claims 1-12, wherein the detecting step comprises mass spectrometry.

14. The method of any of claims 1-6 and 8-12, wherein the detection is effected by detecting a signal moiety selected from the group consisting of radioisotopes, enzymes, antigens, antibodies, spectrophotometric reagents, chemiluminescent reagents and fluorescent reagents.

15. The method of claim 11 or claim 12, wherein the nucleic acid primer is extended in the presence of at least one dideoxynucleotide.

16. The method of claim 15 or claim 16, wherein the dideoxynucleotide is dideoxyguanosine (ddG).

17. The method of claim 11, wherein the primer is extended in the presence at least two dideoxynucleotides and the dideoxynucleotides are dideoxyguanosine (ddG) and dideoxycytosine (ddC).

18. The method of claim 12, wherein the primer is extended in the presence of at least two dideoxynucleotides and the dideoxy-nucleotides are dideoxyguanosine (ddG) and dideoxycytosine (ddC).

19. A method for indicating a predisposition to cardiovascular disease in a subject, comprising:

the step of detecting in a target nucleic acid obtained from the subject the presence or absence of at least one allelic variant of polymorphic regions of a cytochrome C oxidase subunit VIb (COX6B) gene associated with high serum cholesterol or at least one allelic variant

of polymorphic regions of a N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene associated with low serum HDL wherein the presence of an allelic variant is indicative of a predisposition to cardiovascular disease compared to a subject who does not comprise the allelic variant.

20. The method of claim 19, wherein the allelic variant is of a polymorphic region of the cytochrome C oxidase subunit VIb (COX6B) gene.

21. The method of claim 19, wherein the allelic variant is of a polymorphic region of the N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene.

22. The method of claim 20 or claim 21, wherein the polymorphic region is a single nucleotide polymorphism (SNP).

23. The method of claim 22, wherein the SNP is at position 86 of the cytochrome C oxidase subunit VIb (COX6B) gene coding sequence and the allelic variant is represented by a T nucleotide in the sense strand or an A nucleotide in the corresponding position in the antisense strand.

24. The method of claim 22, wherein the SNP is at position 2577 of the N-acetylgluosaminyl transferase component GPI-1 (GPI-1) gene sequence and the allelic variant is represented by an A nucleotide in the sense strand or a T nucleotide in the corresponding position in the antisense strand.

25. The method of claim 19, wherein the detecting step is by a method selected from the group consisting of allele specific hybridization, primer specific extension, oligonucleotide ligation assay, restriction enzyme site analysis and single-stranded conformation polymorphism analysis.

26. The method of claim 23, further comprising:

(a) hybridizing a target nucleic acid comprising a cytochrome C oxidase subunit VIb (COX6B)-encoding nucleic acid or fragment thereof with a nucleic acid primer that hybridizes adjacent to nucleotide 86 of the coding sequence of the COX6B gene;

(b) extending the nucleic acid primer using the target nucleic acid as a template; and

(c) determining the mass of the extended primer to identify the nucleotide present at position 86, thereby determining the presence or absence of the allelic variant.

27. The method of claim 24, further comprising:

(a) hybridizing a target nucleic acid comprising a N-acetylglucosaminyl transferase component GPI-1 (GPI-1)-encoding nucleic acid or fragment thereof with a nucleic acid primer that hybridizes adjacent to nucleotide 2577 of the GPI-1 gene;

(b) extending the nucleic acid primer using the target nucleic acid as a template; and

(c) determining the mass of the extended primer to identify the nucleotide present at position 2577, thereby determining the presence or absence of the allelic variant.

28. The method of claim 19, wherein the detecting step comprises mass spectrometry.

29. The method of claim 19, wherein the detection is effected by detecting a signal moiety selected from the group consisting of: radioisotopes, enzymes, antigens, antibodies, spectrophotometric reagents, chemiluminescent reagents, fluorescent reagents and other light producing reagents.

30. The method of claim 19, further comprising detecting the presence or absence of at least one allelic variant of polymorphic regions of another gene associated with cardiovascular disease, wherein the

presence of the two allelic variants is associated with a predisposition to cardiovascular disease compared to a subject who does not comprise the combination of allelic variants.

31. The method of claim 30, wherein the other gene is selected from the group consisting of cholesterol ester transfer protein, plasma (CETP); apolipoprotein A-IV (APO A4); apolipoprotein A-I (APO A1); apolipoprotein E (APO E); apolipoprotein B (APO B); apolipoprotein C-III (APO C3); a gene encoding lipoprotein lipase (LPL); ATP-binding cassette transporter (ABC 1); paraoxonase 1 (PON 1); paraoxonase 2 (PON 2); 5,10-methylenetetrahydrofolate r reductase (MTHFR); a gene encoding hepatic lipase, E-selectin, G protein beta 3 subunit and angiotensin II type 1 receptor gene.

32. The method of claim 30, wherein the two allelic variants are of the cytochrome C oxidase subunit VIb (COX6B) gene and the N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene.

33. A kit comprising:

(a) at least one probe specific for a polymorphic region of a human gene selected from the group consisting of cytochrome C oxidase subunit VIb (COX6B); N-acetylglucosaminyl transferase component GPI-1 (GPI-1); cholesterol ester transfer protein, plasma (CETP); apolipoprotein A-IV (APO A4); apolipoprotein A-I (APO A1); apolipoprotein E (APO E); apolipoprotein B (APO B); apolipoprotein C-III (APO C3); a gene encoding lipoprotein lipase (LPL); ATP-binding cassette transporter (ABC 1); paraoxonase 1 (PON 1); paraoxonase 2 (PON 2); 5,10-methylenetetrahydrofolate r reductase (MTHFR); a gene encoding hepatic lipase, E-selectin, G protein beta 3 subunit and angiotensin II type 1 receptor gene; and

(b) instructions for use.

34. A method of screening for biologically active agents that modulate serum cholesterol, comprising:

(a) combining a candidate agent with a cell comprising a nucleotide sequence encoding an allelic variant of a cytochrome C oxidase subunit VIb (COX6B) gene associated with high levels of serum cholesterol and operably linked to a promoter such that the nucleotide sequence is expressed as a COX6B protein in the cell; and

(b) determining the affect of the agent upon the expression and/or activity of the COX6B protein.

35. A method of screening for biologically active agents that modulate serum cholesterol, comprising:

(a) combining a candidate agent with a transgenic mouse comprising a transgenic nucleotide sequence stably integrated into the genome of the mouse, wherein the transgenic nucleotide sequence encodes an allelic variant of a cytochrome C oxidase subunit VIb (COX6B) gene associated with high levels of serum cholesterol and operably linked to a promoter, wherein the transgenic nucleotide sequence is expressed and the transgenic animal develops a high level of serum cholesterol; and

(b) determining the affect of the agent upon the serum cholesterol level.

36. The method of claim 34 or claim 37 wherein the allelic variant is at position 86 of the cytochrome C oxidase subunit VIb (COX6B) gene.

37. A method of screening for biologically active agents that modulate serum high density lipoprotein (HDL), comprising:

(a) combining a candidate agent with a cell comprising a nucleotide sequence encoding an allelic variant of a N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene

associated with low levels of serum HDL and operably linked to a promoter such that the nucleotide sequence is expressed as a GPI-1 protein in the cell; and

(b) determining the affect of the agent upon the expression and/or activity of the GPI-1 protein.

38. A method of screening for biologically active agents that modulate serum high density lipoprotein (HDL), comprising:

(a) combining a candidate agent with a transgenic mouse comprising a transgenic nucleotide sequence stably integrated into the genome of the mouse encoding an allelic variant of a N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene associated with low levels of serum HDL operably linked to a promoter, wherein the transgenic nucleotide sequence is expressed and the transgenic animal develops a low level of serum HDL; and

(b) determining the affect of the agent upon the serum HDL level.

39. The method of claim 37 or claim 38, wherein the allelic variant is at position 2577 of the N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene.

40. A method for predicting a response of a subject to a cardiovascular drug, comprising:

detecting the presence or absence of at least one allelic variant of a cytochrome C oxidase subunit VIb (COX6B) gene of the subject associated with high serum cholesterol or at least one allelic variant of a N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene of the subject associated with low serum high density lipoprotein (HDL);

wherein the presence of at least one allelic variant is indicative of a positive response.

41. The method of claim 40, wherein the allelic variant is of the cytochrome C oxidase subunit VIb (COX6B) gene.

42. The method of claim 40, wherein the allelic variant is of the N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene.

43. A method for predicting a response of a subject to a cardiovascular drug, comprising:

detecting the presence or absence of at least one allelic variant of a cytochrome C oxidase subunit VIb (COX6B) gene of the subject associated with high serum cholesterol; and

detecting the presence or absence of or at least one allelic variant of a N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene of the subject associated with low serum high density lipoprotein (HDL);

wherein the presence of at least one allelic variant of the COX6B and at least one allelic variant of the GPI-1 gene is indicative of a positive response.

44. A method for predicting a response of a subject to a biologically active agent that modulates serum cholesterol, comprising:

detecting the presence or absence of at least one allelic variant of a cytochrome C oxidase subunit VIb (COX6B) gene of the subject associated with high cholesterol ;

wherein the presence of at least one allelic variant is indicative of a positive response.

45. A method for predicting a response of a subject to a biologically active agent that modulates serum cholesterol, comprising:

detecting the presence or absence of at least one allelic variant of a cytochrome C oxidase subunit VIb (COX6B) gene of the subject associated with high cholesterol; and

detecting the presence or absence of an allelic variant of at least one other gene of the subject associated with cardiovascular disease, wherein the presence of both allelic variants is indicative of a positive response.

46. The method of claim 44 or claim 45, wherein the allelic variant of the cytochrome C oxidase subunit VIb (COX6B) gene is at position 86.

47. A method for predicting a response of a subject to a biologically active agent that modulates serum high density lipoprotein (HDL), comprising:

detecting the presence or absence of at least one allelic variant of a N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene of the subject associated with low HDL; wherein the presence of an allelic variant is indicative of a positive response.

48. A method for predicting a response of a subject to a biologically active agent that modulates serum high density lipoprotein (HDL) levels, comprising:

(a) detecting the presence or absence of at least one allelic variant of a N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene associated with low HDL of the subject; and

(b) detecting the presence or absence of an allelic variant in at least one other gene of subject associated with cardiovascular disease, wherein the presence of both allelic variants is indicative of a positive response.

49. The method of claim 47 or claim 48, wherein the allelic variant of a N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene is at position 2577.

50. The method of claim 45 or 48, wherein the other gene associated with cardiovascular disease is selected from the group of genes consisting of N-acetylglucosaminyl transferase component GPI (GPI-1) gene, cholesterol ester transfer protein, plasma (CETP); apolipoprotein A-IV (APO A4); apolipoprotein A-I (APO A1); apolipoprotein E (APO E); apolipoprotein B (APO B); apolipoprotein C-III (APO C3); a gene encoding lipoprotein lipase (LPL); ATP-binding cassette transporter

(ABC 1); paraoxonase 1 (PON 1); paraoxonase 2 (PON 2); 5,10-methylenetetrahydrofolate r reductase (MTHFR); a gene encoding hepatic lipase, E-selectin, G protein beta 3 subunit and angiotensin II type 1 receptor gene.

51. A primer or probe that specifically hybridizes adjacent to or at a polymorphic region of a cytochrome C oxidase subunit VIb (COX6B) gene associated with high serum cholesterol in combination with a primer or probe that specifically hybridizes adjacent to or at a polymorphic region of a N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene associated with low HDL.

52. The primers or probes of claim 51, further comprising primers or probes that specifically hybridizes adjacent to or at a polymorphic region of another gene associated with cardiovascular disease.

53. The primers or probes of claim 51, wherein the polymorphic region of the cytochrome C oxidase subunit VIb (COX6B) gene comprises nucleotide 86 of the coding strand and the polymorphic region of the N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene comprises nucleotide 2577.

54. The primers or probes of claim 52, wherein the other gene associated with cardiovascular disease is selected from the group of genes consisting of cholesterol ester transfer protein, plasma (CETP); apolipoprotein A-IV (APO A4); apolipoprotein A-I (APO A1); apolipoprotein E (APO E); apolipoprotein B (APO B); apolipoprotein C-III (APO C3); a gene encoding lipoprotein lipase (LPL); ATP-binding cassette transporter (ABC 1); paraoxonase 1 (PON 1); paraoxonase 2 (PON 2); 5,10-methylenetetrahydrofolate r reductase (MTHFR); a gene encoding hepatic lipase, E-selectin, G protein beta 3 subunit and angiotensin II type 1 receptor gene.

55. A kit for indicating whether a subject has a predisposition to developing cardiovascular disease, comprising:

(a) at least one probe or primer that specifically hybridizes adjacent to or at a polymorphic region of a cytochrome C oxidase subunit VIb (COX6B) gene associated with high serum cholesterol; and

(b) optionally instructions for use.

56. The kit of claim 55, wherein the polymorphic region comprises nucleotide 86 of the coding strand.

57. A kit for indicating whether a subject has a predisposition to developing cardiovascular disease, comprising:

(a) at least one probe or primer that specifically hybridizes adjacent to or at a polymorphic region of a cytochrome C oxidase subunit VIb (COX6B) gene associated with high cholesterol;

(b) at least one probe or primer that specifically hybridizes adjacent to or at a polymorphic region of another gene associated with cardiovascular disease; and

(c) optionally instructions for use.

58. The kit of claim 57, wherein the other gene associated with cardiovascular disease is selected from the group of genes consisting of N-acetylglucosaminyl transferase component GPI-1 (GPI-1); cholesterol ester transfer protein, plasma (CETP); apolipoprotein A-IV (APO A4); apolipoprotein A-I (APO A1); apolipoprotein E (APO E); apolipoprotein B (APO B); apolipoprotein C-III (APO C3); a gene encoding lipoprotein lipase (LPL); ATP-binding cassette transporter (ABC 1); paraoxonase 1 (PON 1); paraoxonase 2 (PON 2); 5,10-methylenetetrahydrofolate r reductase (MTHFR); a gene encoding hepatic lipase, E-selectin, G protein beta 3 subunit and angiotensin II type 1 receptor gene.

59. A kit for indicating whether a subject has a predisposition to developing cardiovascular disease, comprising:

(a) at least one probe or primer that specifically hybridizes adjacent to or at a polymorphic region of a N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene associated with low serum high density lipoprotein (HDL); and

(b) optionally instructions for use.

60. The kit of claim 59, wherein the polymorphic region comprises nucleotide 2577 of the coding strand.

61. A kit for indicating whether a subject has a predisposition to developing cardiovascular disease, comprising:

(a) at least one probe or primer that specifically hybridizes adjacent to or at a polymorphic region of a N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene associated with low serum high density lipoprotein (HDL);

(b) at least one probe or primer that specifically hybridizes adjacent to or at a polymorphic region of another gene associated with cardiovascular disease; and

(c) optionally instructions for use.

62. The kit of claim 61, wherein the other gene associated with cardiovascular disease is selected from the group of genes consisting of cytochrome C oxidase subunit VIb (COX6B); cholesterol ester transfer protein, plasma (CETP); apolipoprotein A-IV (APO A4); apolipoprotein A-I (APO A1); apolipoprotein E (APO E); apolipoprotein B (APO B); apolipoprotein C-III (APO C3); a gene encoding lipoprotein lipase (LPL); ATP-binding cassette transporter (ABC 1); paraoxonase 1 (PON 1); paraoxonase 2 (PON 2); 5,10-methylenetetrahydrofolate r reductase (MTHFR); a gene encoding hepatic lipase, E-selectin, G protein beta 3 subunit and angiotensin II type 1 receptor gene.

63. A kit for indicating whether a subject has a predisposition to developing cardiovascular disease, comprising:

(a) at least one probe or primer that specifically hybridizes adjacent to or at a polymorphic region of a cytochrome C oxidase subunit VIb (COX6B) gene associated with high cholesterol;

(b) at least one probe or primer that specifically hybridizes adjacent to or at a polymorphic region of a N-acetylglucosaminyl transferase component GPI-1 (GP1-1) gene associated with low HDL; and

(c) optionally instructions for use.

64. The kit of claim 63, further comprising at least one probe or primer that specifically hybridizes adjacent to or at a polymorphic region of another gene associated with cardiovascular disease.

65. The kit of claim 64, wherein the other gene associated with cardiovascular disease is selected from the group of genes consisting of cholesterol ester transfer protein, plasma (CETP); apolipoprotein A-IV (APO A4); apolipoprotein A-I (APO A1); apolipoprotein E (APO E); apolipoprotein B (APO B); apolipoprotein C-III (APO C3); a gene encoding lipoprotein lipase (LPL); ATP-binding cassette transporter (ABC 1); paraoxonase 1 (PON 1); paraoxonase 2 (PON 2); 5,10-methylenetetrahydrofolate r reductase (MTHFR); a gene encoding hepatic lipase, E-selectin, G protein beta 3 subunit and angiotensin II type 1 receptor gene.

66. A method of diagnosing a predisposition to cardiovascular disease in a human, said method comprising the steps of:

(a) obtaining a biological sample from the human;

(b) isolating DNA from the biological sample; and

(c) detecting the presence or absence of at least one allelic variant of a cytochrome C oxidase subunit VIb (COX6B) gene in the DNA.

67. The method of claim 66, wherein at least one variant is a C to T transversion at position 86 of the cytochrome C oxidase subunit VIb gene (COX6B) coding region.

68. The method of claim 66, further comprising the step of:

detecting the presence or absence of at least one allelic variant of a second gene associated with cardiovascular disease.

69. The method of claim 68, wherein the second gene is selected from the group consisting of human N-acetylglucosaminyl transferase component GPI-1 (GPI-1); cholesterol ester transfer protein, plasma (CETP); apolipoprotein A-IV (APO A4); apolipoprotein A-I (APO A1); apolipoprotein E (APO E); apolipoprotein B (APO B); apolipoprotein C-III (APO C3); a gene encoding lipoprotein lipase (LPL); ATP-binding cassette transporter (ABC 1); paraoxonase 1 (PON 1); paraoxonase 2 (PON 2); 5,10-methylenetetrahydrofolate r reductase (MTHFR); a gene encoding hepatic lipase, E-selectin, G protein beta 3 subunit and angiotensin II type 1 receptor gene.

70. The method of claim 68, wherein the detecting step is performed by an assay selected from the group consisting of allele specific hybridization, primer specific extension, oligonucleotide ligation, restriction enzyme site analysis, and single-stranded conformation polymorphism analysis.

71. A method of diagnosing a predisposition to cardiovascular disease in a human, said method comprising the steps of:

(a) obtaining a biological sample from the human;

(b) isolating DNA from the biological sample; and

(c) detecting the presence or absence of at least one allelic variant of a N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene in the DNA.

72. The method of claim 71, wherein the detecting step is performed by an assay selected from the group consisting of allele

specific hybridization, primer specific extension, oligonucleotide ligation, restriction enzyme site analysis, and single-stranded conformation polymorphism analysis.

73. The method of claim 71, wherein at least one variant is a G to A transversion at position 2577 of a N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene.

74. A method of determining a response of a human to a cardiovascular drug, said method comprising the steps of:

(a) obtaining a biological sample from the human;

(b) isolating DNA from the biological sample; and

(c) detecting the presence or absence of at least one allelic variant of a cytochrome C oxidase subunit VIb (COX6B) gene in the DNA or at least one allelic variant of a N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene in the DNA.

75. The method of claim 74, wherein the detecting step is performed by an assay selected from the group consisting of allele specific hybridization, primer specific extension, oligonucleotide ligation, restriction enzyme site analysis, and single-stranded conformation polymorphism analysis.

76. A microarray, comprising:

an isolated nucleic acid molecule comprising a sequence of nucleotides of a polymorphic region from a human cytochrome C oxidase subunit VIb (COX6B) gene linked to a solid support.

77. The microarray of claim 76, wherein the polymorphic region comprises position 86 of the human cytochrome C oxidase subunit VIb (COX6B) coding region.

78. A microarray, comprising:

an isolated nucleic acid molecule comprising a sequence of nucleotides uence of a polymorphic region from a human N-

acetylglucosaminyl transferase component GPI-1 (GPI-1) gene linked to a solid support.

79. The microarray of claim 78, wherein the polymorphic region comprises a locus selected from the group consisting of position 2577 of the human N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene, position 2829 of the human GPI-1 gene, position 2519 of the human GPI-1 gene, position 2289 of the human GPI-1 gene, position 1938 of the human GPI-1 gene, position 1563 of the human GPI-1 gene, position 2656 of the human GPI-1 gene, and position 2664 of the human GPI-1 gene.

80. The microarray of claim 91, wherein the polymorphic region comprises position 2577 of the human N-acetylglucosaminyl transferase component GPI-1 (GPI-1) gene.

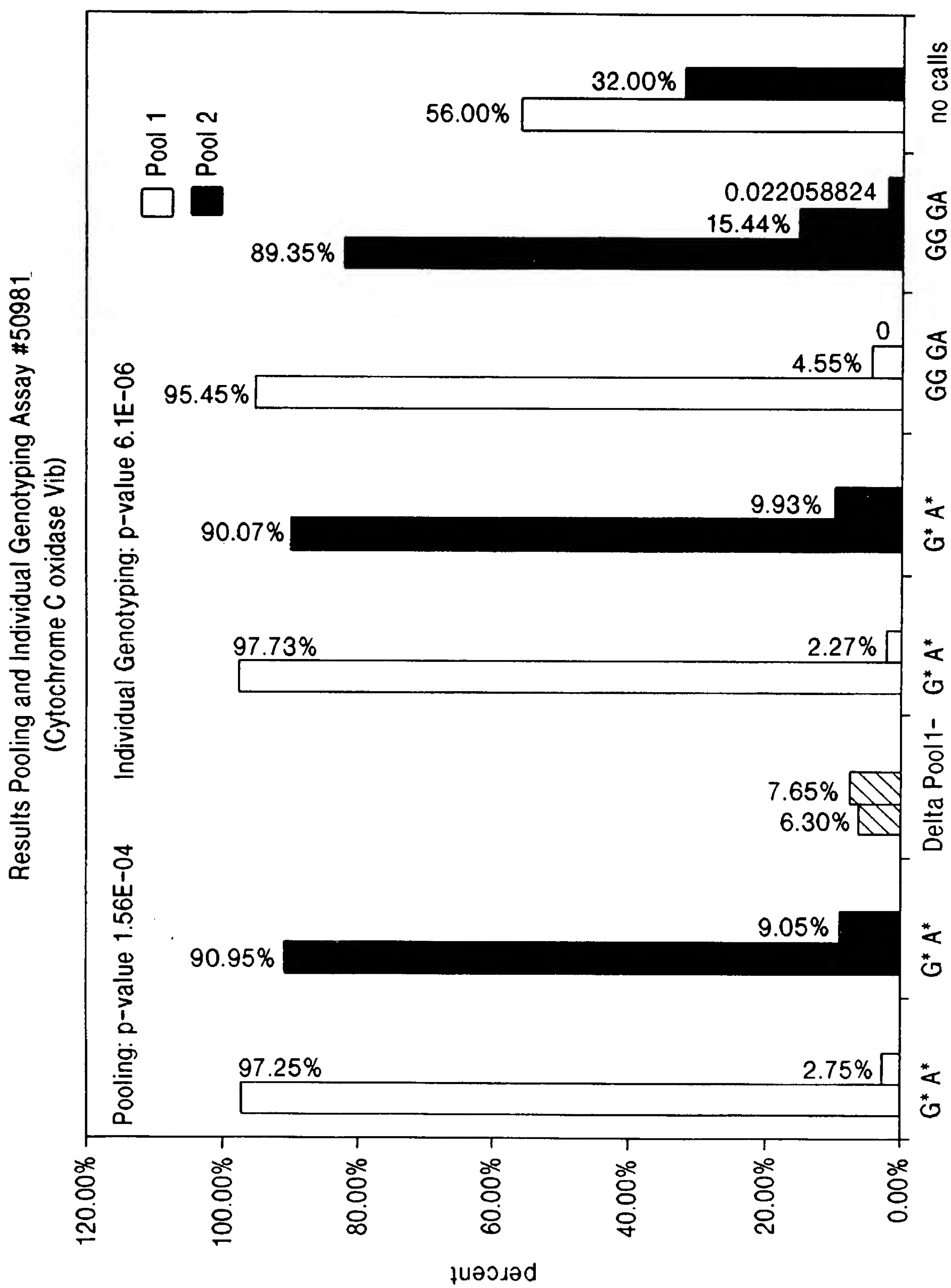
Results Pooling and Individual Genotyping Assay #50981
(Cytochrome C oxidase VIb)
Pooling: p-value 1.56E-04
Individual Genotyping: p-value 6.1E-06
Pool 1
Pool 2
percent
120.00%
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
97.25%
2.75%
G* A*
90.95%
9.05%
G* A*
6.30%
7.65%
Delta Pool1-Pool2
97.73%
2.27%
G* A*
90.07%
9.93%
G* A*
95.45%
4.55%
0
GG GA AA
89.35%
15.44%
0.022058824
GG GA AA
56.00%
32.00%
no calls

FIG. 1

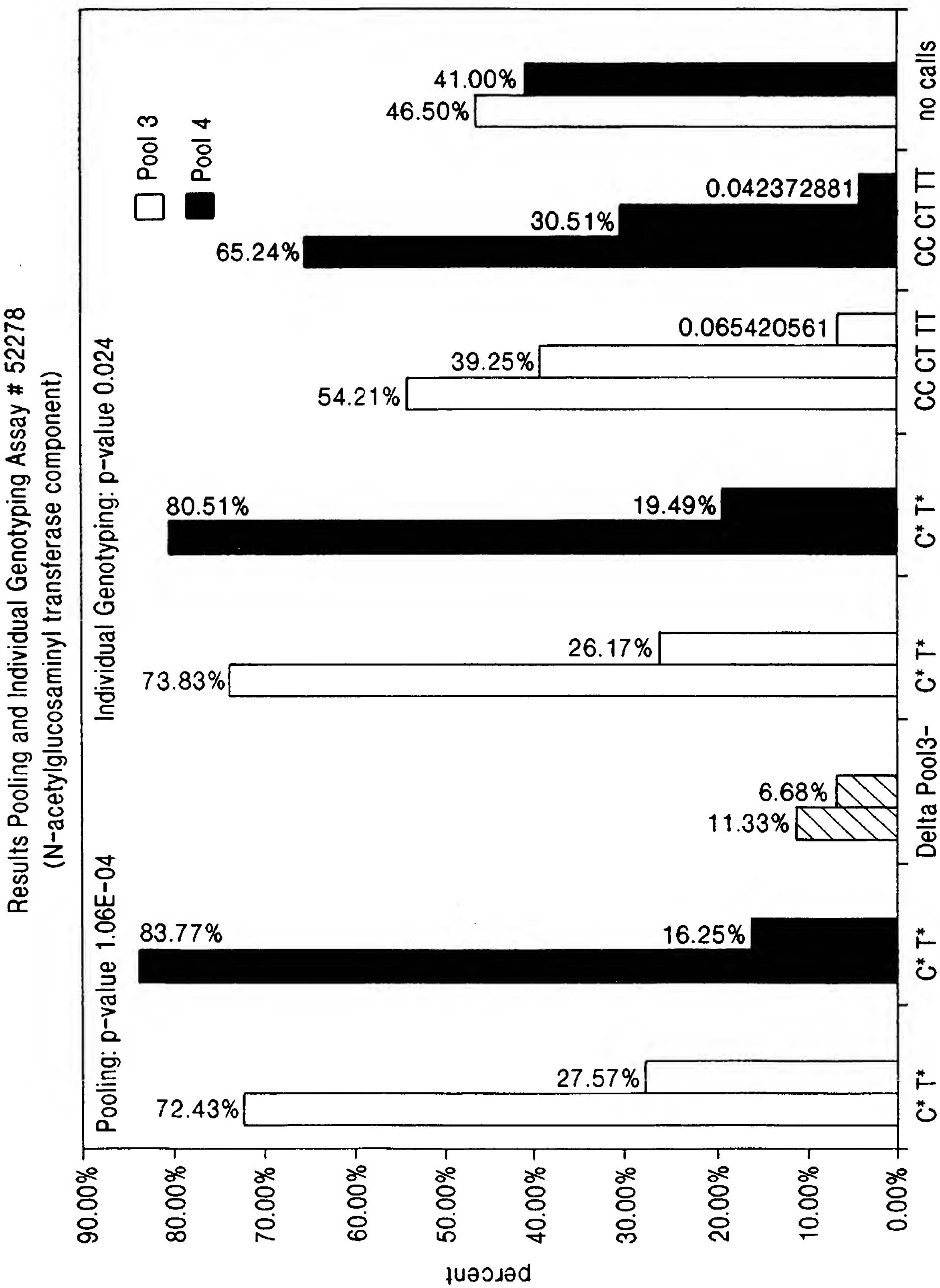
Results Pooling and Individual Genotyping Assay # 52278
(N-acetylglucosaminyl transferase component)
Pooling: p-value 1.06E-04
Individual Genotyping: p-value 0.024
Pool 3
Pool 4
72.43%
27.57%
83.77%
16.25%
11.33%
6.68%
73.83%
26.17%
80.51%
19.49%
54.21%
39.25%
0.065420561
65.24%
30.51%
0.042372881
46.50%
41.00%
C* T*
C* T*
Delta Pool3-Pool4
C* T*
C* T*
CC CT TT
CC CT TT
no calls
percent
90.00%
80.00%
70.00%
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%

FIG. 2

SEQUENCE LISTING

<110> Braun, Andreas
Bonsal Aruna
Kleyn Patrick

<120> GENES AND POLYMORPHISMS ASSOCIATED WITH
CARDIOVASCULAR DISEASE AND THEIR USE

<130> 24736-2048PC

<140> Unassigned
<141> Herewith

<140> 09/802,640
<141> 03-09-2001

<160> 122

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 439
<212> DNA
<213> Homo Sapien

<220>
<221> CDS
<222> (45)...(305)

<400> 1
ttgagctgca ggttgaatcc ggggtgcctt taggattcag cacc atg gcg gaa gac 56
Met Ala Glu Asp
1

atg gag acc aaa atc aag aac tac aag acc gcc cct ttt gac agc cgc 104
Met Glu Thr Lys Ile Lys Asn Tyr Lys Thr Ala Pro Phe Asp Ser Arg
5 10 15 20

ttc ccc aac cag aac cag act aga aac tgc tgg cag aac tac ctg gac 152
Phe Pro Asn Gln Asn Gln Thr Arg Asn Cys Trp Gln Asn Tyr Leu Asp
25 30 35

ttc cac cgc tgt cag aag gca atg acc gct aaa gga ggc gat atc tct 200
Phe His Arg Cys Gln Lys Ala Met Thr Ala Lys Gly Gly Asp Ile Ser
40 45 50

gtg tgc gaa tgg tac cag cgt gtg tac cag tcc ctc tgc ccc aca tcc 248
Val Cys Glu Trp Tyr Gln Arg Val Tyr Gln Ser Leu Cys Pro Thr Ser
55 60 65

tgg gtc aca gac tgg gat gag caa cgg gct gaa ggc acg ttt ccc ggg 296
Trp Val Thr Asp Trp Asp Glu Gln Arg Ala Glu Gly Thr Phe Pro Gly
70 75 80

aag atc tga actggctgca tctccctttc ctctgtcctc catccttctc 345
Lys Ile *

85

ccaggatggt gaagggggac ctggtaccca gtgatcccca ccccaggatc ctaaatcatg 405
acttacctgc taataaaaac tcattggaaa agtg 439

<210> 2
<211> 86
<212> PRT
<213> Homo Sapien

<400> 2
Met Ala Glu Asp Met Glu Thr Lys Ile Lys Asn Tyr Lys Thr Ala Pro
1 5 10 15
Phe Asp Ser Arg Phe Pro Asn Gln Asn Gln Thr Arg Asn Cys Trp Gln
20 25 30
Asn Tyr Leu Asp Phe His Arg Cys Gln Lys Ala Met Thr Ala Lys Gly
35 40 45
Gly Asp Ile Ser Val Cys Glu Trp Tyr Gln Arg Val Tyr Gln Ser Leu
50 55 60
Cys Pro Thr Ser Trp Val Thr Asp Trp Asp Glu Gln Arg Ala Glu Gly
65 70 75 80
Thr Phe Pro Gly Lys Ile
85

<210> 3
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 3
aggattcagc accatggc 18

<210> 4
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 4
agcggataac aatttcacac aggtagtctg gttctggttg ggg 43

<210> 5
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> MassExtend primer

<400> 5
aatcaagaac tacaagac 18

<210> 6

<211> 2921
<212> DNA
<213> Homo Sapien

<220>
<221> CDS
<222> (103)...(1848)

<400> 6
cagcgagcgc cgtcgtctgc ccgggcccgc ccatcggggt ccccaacccc atccggaccc 60
cgccgcccga gcgcgcggcc ccggaagcac ccgcctcccg gc atg gtg ctc aag 114
Met Val Leu Lys
1

gcc ttc ttc ccc acg tgc tgc gtc tcg gcg gac agc ggg ctg ctg gtg 162
Ala Phe Phe Pro Thr Cys Cys Val Ser Ala Asp Ser Gly Leu Leu Val
5 10 15 20

gga cgg tgg gtg ccg gag cag agc agc gcc gtg gtc ctg gcg gtc ctg 210
Gly Arg Trp Val Pro Glu Gln Ser Ser Ala Val Val Leu Ala Val Leu
25 30 35

cac ttt ccc ttc atc ccc atc cag gtc aag cag ctc ctg gcc cag gtg 258
His Phe Pro Phe Ile Pro Ile Gln Val Lys Gln Leu Leu Ala Gln Val
40 45 50

cgg cag gcc agc cag gtg ggc gtg gcc gtg ctg ggc acc tgg tgc cac 306
Arg Gln Ala Ser Gln Val Gly Val Ala Val Leu Gly Thr Trp Cys His
55 60 65

tgc cgg cag gag ccc gag gag agc ctg ggc cgc ttc ctg gag agc ctg 354
Cys Arg Gln Glu Pro Glu Glu Ser Leu Gly Arg Phe Leu Glu Ser Leu
70 75 80

ggt gct gtc ttc ccc cat gag ccc tgg ctg cgg ctg tgc cgg gag aga 402
Gly Ala Val Phe Pro His Glu Pro Trp Leu Arg Leu Cys Arg Glu Arg
85 90 95 100

ggc ggc acg ttc tgg agc tgc gag gcc acc cac cgg caa gcg ccc act 450
Gly Gly Thr Phe Trp Ser Cys Glu Ala Thr His Arg Gln Ala Pro Thr
105 110 115

gcc ccc ggt gcc cct ggt gag gac cag gtc atg ctc atc ttc tat gac 498
Ala Pro Gly Ala Pro Gly Glu Asp Gln Val Met Leu Ile Phe Tyr Asp
120 125 130

cag cgc cag gtg ttg ctg tca cag cta cac ctg ccc acc gtc ctg ccc 546
Gln Arg Gln Val Leu Leu Ser Gln Leu His Leu Pro Thr Val Leu Pro
135 140 145

gac cgc cag gct gga gcc acc act gcc agc acg ggg ggc ctg gct gcc 594
Asp Arg Gln Ala Gly Ala Thr Thr Ala Ser Thr Gly Gly Leu Ala Ala
150 155 160

gtc ttc gac acg gta gca cgc agt gag gtg ctc ttc cgc agt gac cgc 642
Val Phe Asp Thr Val Ala Arg Ser Glu Val Leu Phe Arg Ser Asp Arg
165 170 175 180

ttt gat gag ggc ccc gtg cgg ctg agc cac tgg cag tcg gag ggc gtg 690
Phe Asp Glu Gly Pro Val Arg Leu Ser His Trp Gln Ser Glu Gly Val
185 190 195

gag gcc agc atc ctc gcg gag ctg gcc agg cga gcc tcg gga ccc att 738
Glu Ala Ser Ile Leu Ala Glu Leu Ala Arg Arg Ala Ser Gly Pro Ile
200 205 210

tgt ctg ctg ttg gcc agc ctg ctg tcg ctg gtc tca gct gtc agt gcc 786
Cys Leu Leu Leu Ala Ser Leu Leu Ser Leu Val Ser Ala Val Ser Ala
215 220 225

tgc cga gtg ttc aag ctc tgg ccc ctg tcc ttc ctc ggg agc aaa ctc 834
Cys Arg Val Phe Lys Leu Trp Pro Leu Ser Phe Leu Gly Ser Lys Leu
230 235 240

tcc acg tgc gaa cag ctc cgg cac cgg ctg gag cac ctc acg cta atc 882
Ser Thr Cys Glu Gln Leu Arg His Arg Leu Glu His Leu Thr Leu Ile
245 250 255 260

ttc agt aca cgg aag gcg gag aac cct gcc cag ctg atg agg aag gcc 930
Phe Ser Thr Arg Lys Ala Glu Asn Pro Ala Gln Leu Met Arg Lys Ala
265 270 275

aac acg gtg gcc tct gtg ctg ctg gac gtg gcc ctg ggc ctc atg ctg 978
Asn Thr Val Ala Ser Val Leu Leu Asp Val Ala Leu Gly Leu Met Leu
280 285 290

ctg tcc tgg ctc cac ggg aga agc cgc atc ggg cat ctg gcc gac gcc 1026
Leu Ser Trp Leu His Gly Arg Ser Arg Ile Gly His Leu Ala Asp Ala
295 300 305

ctc gtt cct gtg gct gac cac gtg gcc gag gag ctc cag cat ctg ctg 1074
Leu Val Pro Val Ala Asp His Val Ala Glu Glu Leu Gln His Leu Leu
310 315 320

cag tgg ctg atg ggt gct ccc gcc ggg ctc aag atg aac cgt gca ctg 1122
Gln Trp Leu Met Gly Ala Pro Ala Gly Leu Lys Met Asn Arg Ala Leu
325 330 335 340

gac cag gtg ctg ggc cgc ttc ttc ctc tac cac atc cac ctg tgg atc 1170
Asp Gln Val Leu Gly Arg Phe Phe Leu Tyr His Ile His Leu Trp Ile
345 350 355

agc tac atc cac ctc atg tcc ccc ttc gtg gag cac atc ctt tgg cac 1218
Ser Tyr Ile His Leu Met Ser Pro Phe Val Glu His Ile Leu Trp His
360 365 370

gtg ggc ctc tcg gcc tgc ctg ggc ctg acg gtg gcc ctg tcc ctc ctc 1266
Val Gly Leu Ser Ala Cys Leu Gly Leu Thr Val Ala Leu Ser Leu Leu
375 380 385

tcg gac att atc gcc ctc ctc acc ttc cac atc tac tgc ttt tac gtc 1314
Ser Asp Ile Ile Ala Leu Leu Thr Phe His Ile Tyr Cys Phe Tyr Val
390 395 400

tat gga gcc agg ctg tac tgc ctg aag atc cat ggc ctg tcc tca ctg 1362
Tyr Gly Ala Arg Leu Tyr Cys Leu Lys Ile His Gly Leu Ser Ser Leu

```
405                 410                 415                 420

tgg cgt ctg ttc cgg ggg aag aag tgg aac gtt ctg cgc cag cgc gtg     1410
Trp Arg Leu Phe Arg Gly Lys Lys Trp Asn Val Leu Arg Gln Arg Val
                425                 430                 435

gac tcc tgt tcc tat gac ctg gac cag ctg ttc atc ggg act ctg ctc     1458
Asp Ser Cys Ser Tyr Asp Leu Asp Gln Leu Phe Ile Gly Thr Leu Leu
            440                 445                 450

ttc acc atc ctg ctc ttc ctc ctg cct acc aca gcc ctg tac tac ctg     1506
Phe Thr Ile Leu Leu Phe Leu Leu Pro Thr Thr Ala Leu Tyr Tyr Leu
        455                 460                 465

gtg ttc acc ctg ctc cgg ctc ctg gtg gtc gcc gtg cag ggc ctg atc     1554
Val Phe Thr Leu Leu Arg Leu Leu Val Val Ala Val Gln Gly Leu Ile
    470                 475                 480

cat ctg ctg gtg gac ctc atc aac tcc ctg ccg ctg tac tca ctg ggt     1602
His Leu Leu Val Asp Leu Ile Asn Ser Leu Pro Leu Tyr Ser Leu Gly
485                 490                 495                 500

ctt cgg ctc tgc cgg ccc tac agg ctg gcg gct ggc gtg aag ttc cgt     1650
Leu Arg Leu Cys Arg Pro Tyr Arg Leu Ala Ala Gly Val Lys Phe Arg
                505                 510                 515

gtc ctc cgg cac gag gcc agc agg ccc ctc cgc ctc ctg atg cag ata     1698
Val Leu Arg His Glu Ala Ser Arg Pro Leu Arg Leu Leu Met Gln Ile
            520                 525                 530

aac cca ctg ccc tac agc cgc gtg gtg cac acc tac cgc ctc ccc agc     1746
Asn Pro Leu Pro Tyr Ser Arg Val Val His Thr Tyr Arg Leu Pro Ser
        535                 540                 545

tgt ggc tgc cac ccc aag cac tcc tgg ggc gcc ctg tgc cgc aag ctg     1794
Cys Gly Cys His Pro Lys His Ser Trp Gly Ala Leu Cys Arg Lys Leu
    550                 555                 560

ttc ctt ggg gag ctc atc tac ccc tgg agg cag aga ggg gac aag cag     1842
Phe Leu Gly Glu Leu Ile Tyr Pro Trp Arg Gln Arg Gly Asp Lys Gln
565                 570                 575                 580

gac tga gggaactgct ggctcgcctg gcaccaccac acggccacag ccagccatct      1898
Asp  *

gctctgccag ggtggcacca gctcagctgg cgcatgtccc gtgctttgtg gacgctgctg   1958
tgtgctcctg aacacggcag gccctgctat cacaccttgg gcttggaggt cattgggagt   2018
gagcagatgt gggggtggcc agccaggctg gccgcactcc atcactggca ctgcctgcct   2078
tgggacccgc ttcccacctg ctgcggtcac catggtggcg agcacagcaa ccccaggtgt   2138
ccagagcact gccccatgcc caccctgcat acccaggtcc agagggtccg tccaccacag   2198
cagccccagg tggagggctg gtctccctgg gggctcccca gtggctctgc cctggctgtg   2258
ggggtggagg gaccttgcca ggatgaaccc tccagtccca ggcaccctct agctccctca   2318
gccgaacagc accctgcatc tgggggattg aagcagtcgc tgacccccgt ccccagcggg   2378
cccgggccct cactccctga accacacggg gtttatttgc ggatgttccc tggagaggtc   2438
gctttgtgaa gaaaccatca gcaggctgtg agcatcgcca ggctgctgtg gggcgggag    2498
cagcctcagt gtcaagggcc tgcccactga cccagccgta cctattcgtc cacggtgccc   2558
cgtagcagca ggtcctgcgg ccaaatctgt ctcccttcat gggcctccca gggaaggagg   2618
```

aagccctgct gtgcagacac ctctgtggcc ccccaggggt gtgagcggcc tggggagggg 2678
gccgtggcac tgaggccgaa agtgcctgcc agacggcacg gtctgggtgc gggtgttccc 2738
tgtgagcccg agtccgcttc aggaggggag cctgcaggtg ccggctggtg aggggatgac 2798
gcgctgtggg tgggaggagg cagcgcccat ctcagcagca ccaggactgc ctgggactcc 2858
ctggcaaccc agcaccgggg aagccgtcag ctgctgtgac aataaaacct gccccgtgtc 2918
tgg 2921

<210> 7
<211> 581
<212> PRT
<213> Homo Sapien

<400> 7
Met Val Leu Lys Ala Phe Phe Pro Thr Cys Cys Val Ser Ala Asp Ser
1 5 10 15
Gly Leu Leu Val Gly Arg Trp Val Pro Glu Gln Ser Ser Ala Val Val
20 25 30
Leu Ala Val Leu His Phe Pro Phe Ile Pro Ile Gln Val Lys Gln Leu
35 40 45
Leu Ala Gln Val Arg Gln Ala Ser Gln Val Gly Val Ala Val Leu Gly
50 55 60
Thr Trp Cys His Cys Arg Gln Glu Pro Glu Glu Ser Leu Gly Arg Phe
65 70 75 80
Leu Glu Ser Leu Gly Ala Val Phe Pro His Glu Pro Trp Leu Arg Leu
85 90 95
Cys Arg Glu Arg Gly Gly Thr Phe Trp Ser Cys Glu Ala Thr His Arg
100 105 110
Gln Ala Pro Thr Ala Pro Gly Ala Pro Gly Glu Asp Gln Val Met Leu
115 120 125
Ile Phe Tyr Asp Gln Arg Gln Val Leu Leu Ser Gln Leu His Leu Pro
130 135 140
Thr Val Leu Pro Asp Arg Gln Ala Gly Ala Thr Thr Ala Ser Thr Gly
145 150 155 160
Gly Leu Ala Ala Val Phe Asp Thr Val Ala Arg Ser Glu Val Leu Phe
165 170 175
Arg Ser Asp Arg Phe Asp Glu Gly Pro Val Arg Leu Ser His Trp Gln
180 185 190
Ser Glu Gly Val Glu Ala Ser Ile Leu Ala Glu Leu Ala Arg Arg Ala
195 200 205
Ser Gly Pro Ile Cys Leu Leu Leu Ala Ser Leu Leu Ser Leu Val Ser
210 215 220
Ala Val Ser Ala Cys Arg Val Phe Lys Leu Trp Pro Leu Ser Phe Leu
225 230 235 240
Gly Ser Lys Leu Ser Thr Cys Glu Gln Leu Arg His Arg Leu Glu His
245 250 255
Leu Thr Leu Ile Phe Ser Thr Arg Lys Ala Glu Asn Pro Ala Gln Leu
260 265 270
Met Arg Lys Ala Asn Thr Val Ala Ser Val Leu Leu Asp Val Ala Leu
275 280 285
Gly Leu Met Leu Leu Ser Trp Leu His Gly Arg Ser Arg Ile Gly His
290 295 300
Leu Ala Asp Ala Leu Val Pro Val Ala Asp His Val Ala Glu Glu Leu
305 310 315 320
Gln His Leu Leu Gln Trp Leu Met Gly Ala Pro Ala Gly Leu Lys Met
325 330 335
Asn Arg Ala Leu Asp Gln Val Leu Gly Arg Phe Phe Leu Tyr His Ile
340 345 350
His Leu Trp Ile Ser Tyr Ile His Leu Met Ser Pro Phe Val Glu His

```
        355                 360                 365
Ile Leu Trp His Val Gly Leu Ser Ala Cys Leu Gly Leu Thr Val Ala
    370                 375                 380
Leu Ser Leu Leu Ser Asp Ile Ile Ala Leu Leu Thr Phe His Ile Tyr
385                 390                 395                 400
Cys Phe Tyr Val Tyr Gly Ala Arg Leu Tyr Cys Leu Lys Ile His Gly
                405                 410                 415
Leu Ser Ser Leu Trp Arg Leu Phe Arg Gly Lys Lys Trp Asn Val Leu
            420                 425                 430
Arg Gln Arg Val Asp Ser Cys Ser Tyr Asp Leu Asp Gln Leu Phe Ile
        435                 440                 445
Gly Thr Leu Leu Phe Thr Ile Leu Leu Phe Leu Leu Pro Thr Thr Ala
    450                 455                 460
Leu Tyr Tyr Leu Val Phe Thr Leu Leu Arg Leu Leu Val Val Ala Val
465                 470                 475                 480
Gln Gly Leu Ile His Leu Leu Val Asp Leu Ile Asn Ser Leu Pro Leu
                485                 490                 495
Tyr Ser Leu Gly Leu Arg Leu Cys Arg Pro Tyr Arg Leu Ala Ala Gly
            500                 505                 510
Val Lys Phe Arg Val Leu Arg His Glu Ala Ser Arg Pro Leu Arg Leu
        515                 520                 525
Leu Met Gln Ile Asn Pro Leu Pro Tyr Ser Arg Val Val His Thr Tyr
    530                 535                 540
Arg Leu Pro Ser Cys Gly Cys His Pro Lys His Ser Trp Gly Ala Leu
545                 550                 555                 560
Cys Arg Lys Leu Phe Leu Gly Glu Leu Ile Tyr Pro Trp Arg Gln Arg
                565                 570                 575
Gly Asp Lys Gln Asp
            580
```

<210> 8
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 8
agcagggctt cctccttc 18

<210> 9
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 9
agcggataac aatttcacac aggtgaccca gccgtaccta ttc 43

<210> 10
<211> 18
<212> DNA
<213> Artificial Sequence

<220>

<223> MassExtend primer

<400> 10
aagggagaca gatttggc 18

<210> 11
<211> 1790
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (131)...(1612)
<223> Nucleotide sequence encoding Cholesterol ester
transfer protein (CETP)

<400> 11
gtgaatctct ggggccagga agaccctgct gcccggaaga gcctcatgtt ccgtgggggc 60
tgggcggaca tacatatacg ggctccaggc tgaacggctc gggccactta cacaccactg 120
cctgataacc atg ctg gct gcc aca gtc ctg acc ctg gcc ctg ctg ggc 169
Met Leu Ala Ala Thr Val Leu Thr Leu Ala Leu Leu Gly
1 5 10

aat gcc cat gcc tgc tcc aaa ggc acc tcg cac gag gca ggc atc gtg 217
Asn Ala His Ala Cys Ser Lys Gly Thr Ser His Glu Ala Gly Ile Val
15 20 25

tgc cgc atc acc aag cct gcc ctc ctg gtg ttg aac cac gag act gcc 265
Cys Arg Ile Thr Lys Pro Ala Leu Leu Val Leu Asn His Glu Thr Ala
30 35 40 45

aag gtg atc cag acc gcc ttc cag cga gcc agc tac cca gat atc acg 313
Lys Val Ile Gln Thr Ala Phe Gln Arg Ala Ser Tyr Pro Asp Ile Thr
50 55 60

ggc gag aag gcc atg atg ctc ctt ggc caa gtc aag tat ggg ttg cac 361
Gly Glu Lys Ala Met Met Leu Leu Gly Gln Val Lys Tyr Gly Leu His
65 70 75

aac atc cag atc agc cac ttg tcc atc gcc agc agc cag gtg gag ctg 409
Asn Ile Gln Ile Ser His Leu Ser Ile Ala Ser Ser Gln Val Glu Leu
80 85 90

gtg gaa gcc aag tcc att gat gtc tcc att cag aac gtg tct gtg gtc 457
Val Glu Ala Lys Ser Ile Asp Val Ser Ile Gln Asn Val Ser Val Val
95 100 105

ttc aag ggg acc ctg aag tat ggc tac acc act gcc tgg tgg ctg ggt 505
Phe Lys Gly Thr Leu Lys Tyr Gly Tyr Thr Thr Ala Trp Trp Leu Gly
110 115 120 125

att gat cag tcc att gac ttc gag atc gac tct gcc att gac ctc cag 553
Ile Asp Gln Ser Ile Asp Phe Glu Ile Asp Ser Ala Ile Asp Leu Gln
130 135 140

atc aac aca cag ctg acc tgt gac tct ggt aga gtg cgg acc gat gcc 601
Ile Asn Thr Gln Leu Thr Cys Asp Ser Gly Arg Val Arg Thr Asp Ala
145 150 155

```
cct gac tgc tac ctg tct ttc cat aag ctg ctc ctg cat ctc caa ggg      649
Pro Asp Cys Tyr Leu Ser Phe His Lys Leu Leu Leu His Leu Gln Gly
        160                 165                 170

gag cga gag cct ggg tgg atc aag cag ctg ttc aca aat ttc atc tcc      697
Glu Arg Glu Pro Gly Trp Ile Lys Gln Leu Phe Thr Asn Phe Ile Ser
    175                 180                 185

ttc acc ctg aag ctg gtc ctg aag gga cag atc tgc aaa gag atc aac      745
Phe Thr Leu Lys Leu Val Leu Lys Gly Gln Ile Cys Lys Glu Ile Asn
190                 195                 200                 205

gtc atc tct aac atc atg gcc gat ttt gtc cag aca agg gct gcc agc      793
Val Ile Ser Asn Ile Met Ala Asp Phe Val Gln Thr Arg Ala Ala Ser
                210                 215                 220

atc ctt tca gat gga gac att ggg gtg gac att tcc ctg aca ggt gat      841
Ile Leu Ser Asp Gly Asp Ile Gly Val Asp Ile Ser Leu Thr Gly Asp
            225                 230                 235

ccc gtc atc aca gcc tcc tac ctg gag tcc cat cac aag ggt cat ttc      889
Pro Val Ile Thr Ala Ser Tyr Leu Glu Ser His His Lys Gly His Phe
        240                 245                 250

atc tac aag aat gtc tca gag gac ctc ccc ctc ccc acc ttc tcg ccc      937
Ile Tyr Lys Asn Val Ser Glu Asp Leu Pro Leu Pro Thr Phe Ser Pro
    255                 260                 265

aca ctg ctg ggg gac tcc cgc atg ctg tac ttc tgg ttc tct gag cga      985
Thr Leu Leu Gly Asp Ser Arg Met Leu Tyr Phe Trp Phe Ser Glu Arg
270                 275                 280                 285

gtc ttc cac tcg ctg gcc aag gta gct ttc cag gat ggc cgc ctc atg     1033
Val Phe His Ser Leu Ala Lys Val Ala Phe Gln Asp Gly Arg Leu Met
                290                 295                 300

ctc agc ctg atg gga gac gag ttc aag gca gtg ctg gag acc tgg ggc     1081
Leu Ser Leu Met Gly Asp Glu Phe Lys Ala Val Leu Glu Thr Trp Gly
            305                 310                 315

ttc aac acc aac cag gaa atc ttc caa gag gtt gtc ggc ggc ttc ccc     1129
Phe Asn Thr Asn Gln Glu Ile Phe Gln Glu Val Val Gly Gly Phe Pro
        320                 325                 330

agc cag gcc caa gtc acc gtc cac tgc ctc aag atg ccc aag atc tcc     1177
Ser Gln Ala Gln Val Thr Val His Cys Leu Lys Met Pro Lys Ile Ser
    335                 340                 345

tgc caa aac aag gga gtc gtg gtc aat tct tca gtg atg gtg aaa ttc     1225
Cys Gln Asn Lys Gly Val Val Val Asn Ser Ser Val Met Val Lys Phe
350                 355                 360                 365

ctc ttt cca cgc cca gac cag caa cat tct gta gct tac aca ttt gaa     1273
Leu Phe Pro Arg Pro Asp Gln Gln His Ser Val Ala Tyr Thr Phe Glu
                370                 375                 380

gag gat atc gtg act acc gtc cag gcc tcc tat tct aag aaa aag ctc     1321
```

```
Glu Asp Ile Val Thr Thr Val Gln Ala Ser Tyr Ser Lys Lys Lys Leu
            385                 390                 395

ttc tta agc ctc ttg gat ttc cag att aca cca aag act gtt tcc aac      1369
Phe Leu Ser Leu Leu Asp Phe Gln Ile Thr Pro Lys Thr Val Ser Asn
        400                 405                 410

ttg act gag agc agc tcc gag tcc atc cag agc ttc ctg cag tca atg      1417
Leu Thr Glu Ser Ser Ser Glu Ser Ile Gln Ser Phe Leu Gln Ser Met
    415                 420                 425

atc acc gct gtg ggc atc cct gag gtc atg tct cgg ctc gag gta gtg      1465
Ile Thr Ala Val Gly Ile Pro Glu Val Met Ser Arg Leu Glu Val Val
430                 435                 440                 445

ttt aca gcc ctc atg aac agc aaa ggc gtg agc ctc ttc gac atc atc      1513
Phe Thr Ala Leu Met Asn Ser Lys Gly Val Ser Leu Phe Asp Ile Ile
                450                 455                 460

aac cct gag att atc act cga gat ggc ttc ctg ctg ctg cag atg gac      1561
Asn Pro Glu Ile Ile Thr Arg Asp Gly Phe Leu Leu Leu Gln Met Asp
            465                 470                 475

ttt ggc ttc cct gag cac ctg ctg gtg gat ttc ctc cag agc ttg agc      1609
Phe Gly Phe Pro Glu His Leu Leu Val Asp Phe Leu Gln Ser Leu Ser
        480                 485                 490

tag aagtctccaa ggaggtcggg atggggcttg tagcagaagg caagcaccag           1662
 *


gctcacagct ggaaccctgg tgtctcctcc agcgtggtgg aagttgggtt aggagtacgg    1722
agatggagat tggctcccaa ctcctcccta tcctaaaggc ccactggcat taaagtgctg    1782
tatccaag                                                             1790

<210> 12
<211> 493
<212> PRT
<213> Homo sapien

<400> 12
Met Leu Ala Ala Thr Val Leu Thr Leu Ala Leu Leu Gly Asn Ala His
 1               5                  10                  15
Ala Cys Ser Lys Gly Thr Ser His Glu Ala Gly Ile Val Cys Arg Ile
            20                  25                  30
Thr Lys Pro Ala Leu Leu Val Leu Asn His Glu Thr Ala Lys Val Ile
        35                  40                  45
Gln Thr Ala Phe Gln Arg Ala Ser Tyr Pro Asp Ile Thr Gly Glu Lys
    50                  55                  60
Ala Met Met Leu Leu Gly Gln Val Lys Tyr Gly Leu His Asn Ile Gln
65                  70                  75                  80
Ile Ser His Leu Ser Ile Ala Ser Ser Gln Val Glu Leu Val Glu Ala
                85                  90                  95
Lys Ser Ile Asp Val Ser Ile Gln Asn Val Ser Val Val Phe Lys Gly
            100                 105                 110
Thr Leu Lys Tyr Gly Tyr Thr Thr Ala Trp Trp Leu Gly Ile Asp Gln
        115                 120                 125
Ser Ile Asp Phe Glu Ile Asp Ser Ala Ile Asp Leu Gln Ile Asn Thr
```

```
    130                     135                     140
Gln Leu Thr Cys Asp Ser Gly Arg Val Arg Thr Asp Ala Pro Asp Cys
145                 150                 155                 160
Tyr Leu Ser Phe His Lys Leu Leu Leu His Leu Gln Gly Glu Arg Glu
                165                 170                 175
Pro Gly Trp Ile Lys Gln Leu Phe Thr Asn Phe Ile Ser Phe Thr Leu
            180                 185                 190
Lys Leu Val Leu Lys Gly Gln Ile Cys Lys Glu Ile Asn Val Ile Ser
        195                 200                 205
Asn Ile Met Ala Asp Phe Val Gln Thr Arg Ala Ala Ser Ile Leu Ser
    210                 215                 220
Asp Gly Asp Ile Gly Val Asp Ile Ser Leu Thr Gly Asp Pro Val Ile
225                 230                 235                 240
Thr Ala Ser Tyr Leu Glu Ser His His Lys Gly His Phe Ile Tyr Lys
                245                 250                 255
Asn Val Ser Glu Asp Leu Pro Leu Pro Thr Phe Ser Pro Thr Leu Leu
            260                 265                 270
Gly Asp Ser Arg Met Leu Tyr Phe Trp Phe Ser Glu Arg Val Phe His
        275                 280                 285
Ser Leu Ala Lys Val Ala Phe Gln Asp Gly Arg Leu Met Leu Ser Leu
    290                 295                 300
Met Gly Asp Glu Phe Lys Ala Val Leu Glu Thr Trp Gly Phe Asn Thr
305                 310                 315                 320
Asn Gln Glu Ile Phe Gln Glu Val Val Gly Gly Phe Pro Ser Gln Ala
                325                 330                 335
Gln Val Thr Val His Cys Leu Lys Met Pro Lys Ile Ser Cys Gln Asn
            340                 345                 350
Lys Gly Val Val Val Asn Ser Ser Val Met Val Lys Phe Leu Phe Pro
        355                 360                 365
Arg Pro Asp Gln Gln His Ser Val Ala Tyr Thr Phe Glu Glu Asp Ile
    370                 375                 380
Val Thr Thr Val Gln Ala Ser Tyr Ser Lys Lys Lys Leu Phe Leu Ser
385                 390                 395                 400
Leu Leu Asp Phe Gln Ile Thr Pro Lys Thr Val Ser Asn Leu Thr Glu
                405                 410                 415
Ser Ser Ser Glu Ser Ile Gln Ser Phe Leu Gln Ser Met Ile Thr Ala
            420                 425                 430
Val Gly Ile Pro Glu Val Met Ser Arg Leu Glu Val Val Phe Thr Ala
        435                 440                 445
Leu Met Asn Ser Lys Gly Val Ser Leu Phe Asp Ile Ile Asn Pro Glu
    450                 455                 460
Ile Ile Thr Arg Asp Gly Phe Leu Leu Leu Gln Met Asp Phe Gly Phe
465                 470                 475                 480
Pro Glu His Leu Leu Val Asp Phe Leu Gln Ser Leu Ser
                485                 490
```

<210> 13
<211> 3549
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (175)...(1602)
<223> Nucleotide sequence encoding lipoprotein lipase
(LPL)

<400> 13

```
cccctcttcc tcctcctcaa gggaaagctg cccacttcta gctgccctgc catccccttt      60
aaagggcgac ttgctcagcg ccaaaccgcg gctccagccc tctccagcct ccggctcagc     120
cggctcatca gtcggtccgc gccttgcagc tcctccagag ggacgcgccc cgag atg       177
                                                            Met
                                                             1

gag agc aaa gcc ctg ctc gtg ctg act ctg gcc gtg tgg ctc cag agt       225
Glu Ser Lys Ala Leu Leu Val Leu Thr Leu Ala Val Trp Leu Gln Ser
            5                   10                  15

ctg acc gcc tcc cgc gga ggg gtg gcc gcc gcc gac caa aga aga gat       273
Leu Thr Ala Ser Arg Gly Gly Val Ala Ala Ala Asp Gln Arg Arg Asp
        20                  25                  30

ttt atc gac atc gaa agt aaa ttt gcc cta agg acc cct gaa gac aca       321
Phe Ile Asp Ile Glu Ser Lys Phe Ala Leu Arg Thr Pro Glu Asp Thr
    35                  40                  45

gct gag gac act tgc cac ctc att ccc gga gta gca gag tcc gtg gct       369
Ala Glu Asp Thr Cys His Leu Ile Pro Gly Val Ala Glu Ser Val Ala
50                  55                  60                  65

acc tgt cat ttc aat cac agc agc aaa acc ttc atg gtg atc cat ggc       417
Thr Cys His Phe Asn His Ser Ser Lys Thr Phe Met Val Ile His Gly
                70                  75                  80

tgg acg gta aca gga atg tat gag agt tgg gtg cca aaa ctt gtg gcc       465
Trp Thr Val Thr Gly Met Tyr Glu Ser Trp Val Pro Lys Leu Val Ala
            85                  90                  95

gcc ctg tac aag aga gaa cca gac tcc aat gtc att gtg gtg gac tgg       513
Ala Leu Tyr Lys Arg Glu Pro Asp Ser Asn Val Ile Val Val Asp Trp
        100                 105                 110

ctg tca cgg gct cag gag cat tac cca gtg tcc gcg ggc tac acc aaa       561
Leu Ser Arg Ala Gln Glu His Tyr Pro Val Ser Ala Gly Tyr Thr Lys
    115                 120                 125

ctg gtg gga cag gat gtg gcc cgg ttt atc aac tgg atg gag gag gag       609
Leu Val Gly Gln Asp Val Ala Arg Phe Ile Asn Trp Met Glu Glu Glu
130                 135                 140                 145

ttt aac tac cct ctg gac aat gtc cat ctc ttg gga tac agc ctt gga       657
Phe Asn Tyr Pro Leu Asp Asn Val His Leu Leu Gly Tyr Ser Leu Gly
                150                 155                 160

gcc cat gct gct ggc att gca gga agt ctg acc aat aag aaa gtc aac       705
Ala His Ala Ala Gly Ile Ala Gly Ser Leu Thr Asn Lys Lys Val Asn
            165                 170                 175

aga att act ggc ctc gat cca gct gga cct aac ttt gag tat gca gaa       753
Arg Ile Thr Gly Leu Asp Pro Ala Gly Pro Asn Phe Glu Tyr Ala Glu
        180                 185                 190

gcc ccg agt cgt ctt tct cct gat gat gca gat ttt gta gac gtc tta       801
Ala Pro Ser Arg Leu Ser Pro Asp Asp Ala Asp Phe Val Asp Val Leu
    195                 200                 205
```

```
cac aca ttc acc aga ggg tcc cct ggt cga agc att gga atc cag aaa      849
His Thr Phe Thr Arg Gly Ser Pro Gly Arg Ser Ile Gly Ile Gln Lys
210                 215                 220                 225

cca gtt ggg cat gtt gac att tac ccg aat gga ggt act ttt cag cca      897
Pro Val Gly His Val Asp Ile Tyr Pro Asn Gly Gly Thr Phe Gln Pro
                230                 235                 240

gga tgt aac att gga gaa gct atc cgc gtg att gca gag aga gga ctt      945
Gly Cys Asn Ile Gly Glu Ala Ile Arg Val Ile Ala Glu Arg Gly Leu
            245                 250                 255

gga gat gtg gac cag cta gtg aag tgc tcc cac gag cgc tcc att cat      993
Gly Asp Val Asp Gln Leu Val Lys Cys Ser His Glu Arg Ser Ile His
        260                 265                 270

ctc ttc atc gac tct ctg ttg aat gaa gaa aat cca agt aag gcc tac     1041
Leu Phe Ile Asp Ser Leu Leu Asn Glu Glu Asn Pro Ser Lys Ala Tyr
    275                 280                 285

agg tgc agt tcc aag gaa gcc ttt gag aaa ggg ctc tgc ttg agt tgt     1089
Arg Cys Ser Ser Lys Glu Ala Phe Glu Lys Gly Leu Cys Leu Ser Cys
290                 295                 300                 305

aga aag aac cgc tgc aac aat ctg ggc tat gag atc aat aaa gtc aga     1137
Arg Lys Asn Arg Cys Asn Asn Leu Gly Tyr Glu Ile Asn Lys Val Arg
                310                 315                 320

gcc aaa aga agc agc aaa atg tac ctg aag act cgt tct cag atg ccc     1185
Ala Lys Arg Ser Ser Lys Met Tyr Leu Lys Thr Arg Ser Gln Met Pro
            325                 330                 335

tac aaa gtc ttc cat tac caa gta aag att cat ttt tct ggg act gag     1233
Tyr Lys Val Phe His Tyr Gln Val Lys Ile His Phe Ser Gly Thr Glu
        340                 345                 350

agt gaa acc cat acc aat cag gcc ttt gag att tct ctg tat ggc acc     1281
Ser Glu Thr His Thr Asn Gln Ala Phe Glu Ile Ser Leu Tyr Gly Thr
    355                 360                 365

gtg gcc gag agt gag aac atc cca ttc act ctg cct gaa gtt tcc aca     1329
Val Ala Glu Ser Glu Asn Ile Pro Phe Thr Leu Pro Glu Val Ser Thr
370                 375                 380                 385

aat aag acc tac tcc ttc cta att tac aca gag gta gat att gga gaa     1377
Asn Lys Thr Tyr Ser Phe Leu Ile Tyr Thr Glu Val Asp Ile Gly Glu
                390                 395                 400

cta ctc atg ttg aag ctc aaa tgg aag agt gat tca tac ttt agc tgg     1425
Leu Leu Met Leu Lys Leu Lys Trp Lys Ser Asp Ser Tyr Phe Ser Trp
            405                 410                 415

tca gac tgg tgg agc agt ccc ggc ttc gcc att cag aag atc aga gta     1473
Ser Asp Trp Trp Ser Ser Pro Gly Phe Ala Ile Gln Lys Ile Arg Val
        420                 425                 430

aaa gca gga gag act cag aaa aag gtg atc ttc tgt tct agg gag aaa     1521
Lys Ala Gly Glu Thr Gln Lys Lys Val Ile Phe Cys Ser Arg Glu Lys
```

435 440 445

gtg tct cat ttg cag aaa gga aag gca cct gcg gta ttt gtg aaa tgc 1569
Val Ser His Leu Gln Lys Gly Lys Ala Pro Ala Val Phe Val Lys Cys
450 455 460 465

cat gac aag tct ctg aat aag aag tca ggc tga aactgggcga atctacagaa 1622
His Asp Lys Ser Leu Asn Lys Lys Ser Gly *
470 475

caaagaacgg catgtgaatt ctgtgaagaa tgaagtggag gaagtaactt ttacaaaaca 1682
tacccagtgt ttggggtgtt tcaaaagtgg attttcctga atattaatcc cagccctacc 1742
cttgttagtt attttaggag acagtctcaa gcactaaaaa gtggctaatt caatttatgg 1802
ggtatagtgg ccaaatagca catcctccaa cgttaaaaga cagtggatca tgaaaagtgc 1862
tgttttgtcc tttgagaaag aaataattgt ttgagcgcag agtaaaataa ggctccttca 1922
tgtggcgtat tgggccatag cctataattg gttagaacct cctattttaa ttggaattct 1982
ggatctttcg gactgaggcc ttctcaaact ttactctaag tctccaagaa tacagaaaat 2042
gcttttccgc ggcacgaatc agactcatct acacagcagt atgaatgatg ttttagaatg 2102
attccctctt gctattggaa tgtggtccag acgtcaacca ggaacatgta acttggagag 2162
ggacgaagaa agggtctgat aaacacagag gttttaaaca gtccctacca ttggcctgca 2222
tcatgacaaa gttacaaatt caaggagata taaaatctag atcaattaat tcttaatagg 2282
ctttatcgtt tattgcttaa tccctctctc ccccttcttt tttgtctcaa gattatatta 2342
taataatgtt ctctgggtag gtgttgaaaa tgagcctgta atcctcagct gacacataat 2402
ttgaatggtg cagaaaaaaa aaagataccg taattttatt attagattct ccaaatgatt 2462
ttcatcaatt taaaatcatt caatatctga cagttactct tcagttttag gcttaccttg 2522
gtcatgcttc agttgtactt ccagtgcgtc tcttttgttc ctggctttga catgaaaaga 2582
taggtttgag ttcaaatttt gcattgtgtg agcttctaca gattttagac aaggaccgtt 2642
tttactaagt aaaagggtgg agaggttcct ggggtggatt cctaagcagt gcttgtaaac 2702
catcgcgtgc aatgagccag atggagtacc atgagggttg ttatttgttg tttttaacaa 2762
ctaatcaaga gtgagtgaac aactatttat aaactagatc tcctattttt cagaatgctc 2822
ttctacgtat aaatatgaaa tgataaagat gtcaaatatc tcagaggcta tagctgggaa 2882
cccgactgtg aaagtatgtg atatctgaac acatactaga aagctctgca tgtgtgttgt 2942
ccttcagcat aattcggaag ggaaaacagt cgatcaaggg atgtattgga acatgtcgga 3002
gtagaaattg ttcctgatgt gccagaactt cgaccctttc tctgagagag atgatcgtgc 3062
ctataaatag taggaccaat gttgtgatta acatcatcag gcttggaatg aattctctct 3122
aaaaataaaa tgatgtatga tttgttgttg gcatcccctt tattaattca ttaaatttct 3182
ggatttgggt tgtgacccag ggtgcattaa cttaaaagat tcactaaagc agcacatagc 3242
actgggaact ctggctccga aaaactttgt tatatatatc aaggatgttc tggctttaca 3302
ttttatttat tagctgtaaa tacatgtgtg gatgtgtaaa tggagcttgt acatattgga 3362
aaggtcattg tggctatctg catttataaa tgtgtggtgc taactgtatg tgtctttatc 3422
agtgatggtc tcacagagcc aactcactct tatgaaatgg gctttaacaa aacaagaaag 3482
aaacgtactt aactgtgtga agaaatggaa tcagctttta ataaaattga caacatttta 3542
ttaccac 3549

<210> 14
<211> 475
<212> PRT
<213> Homo sapien

<400> 14
Met Glu Ser Lys Ala Leu Leu Val Leu Thr Leu Ala Val Trp Leu Gln
1 5 10 15
Ser Leu Thr Ala Ser Arg Gly Gly Val Ala Ala Ala Asp Gln Arg Arg
20 25 30
Asp Phe Ile Asp Ile Glu Ser Lys Phe Ala Leu Arg Thr Pro Glu Asp
35 40 45
Thr Ala Glu Asp Thr Cys His Leu Ile Pro Gly Val Ala Glu Ser Val
50 55 60

```
Ala Thr Cys His Phe Asn His Ser Ser Lys Thr Phe Met Val Ile His
65                  70                  75                  80
Gly Trp Thr Val Thr Gly Met Tyr Glu Ser Trp Val Pro Lys Leu Val
                85                  90                  95
Ala Ala Leu Tyr Lys Arg Glu Pro Asp Ser Asn Val Ile Val Val Asp
            100                 105                 110
Trp Leu Ser Arg Ala Gln Glu His Tyr Pro Val Ser Ala Gly Tyr Thr
        115                 120                 125
Lys Leu Val Gly Gln Asp Val Ala Arg Phe Ile Asn Trp Met Glu Glu
    130                 135                 140
Glu Phe Asn Tyr Pro Leu Asp Asn Val His Leu Leu Gly Tyr Ser Leu
145                 150                 155                 160
Gly Ala His Ala Ala Gly Ile Ala Gly Ser Leu Thr Asn Lys Lys Val
                165                 170                 175
Asn Arg Ile Thr Gly Leu Asp Pro Ala Gly Pro Asn Phe Glu Tyr Ala
            180                 185                 190
Glu Ala Pro Ser Arg Leu Ser Pro Asp Asp Ala Asp Phe Val Asp Val
        195                 200                 205
Leu His Thr Phe Thr Arg Gly Ser Pro Gly Arg Ser Ile Gly Ile Gln
    210                 215                 220
Lys Pro Val Gly His Val Asp Ile Tyr Pro Asn Gly Gly Thr Phe Gln
225                 230                 235                 240
Pro Gly Cys Asn Ile Gly Glu Ala Ile Arg Val Ile Ala Glu Arg Gly
                245                 250                 255
Leu Gly Asp Val Asp Gln Leu Val Lys Cys Ser His Glu Arg Ser Ile
            260                 265                 270
His Leu Phe Ile Asp Ser Leu Leu Asn Glu Glu Asn Pro Ser Lys Ala
        275                 280                 285
Tyr Arg Cys Ser Ser Lys Glu Ala Phe Glu Lys Gly Leu Cys Leu Ser
    290                 295                 300
Cys Arg Lys Asn Arg Cys Asn Asn Leu Gly Tyr Glu Ile Asn Lys Val
305                 310                 315                 320
Arg Ala Lys Arg Ser Ser Lys Met Tyr Leu Lys Thr Arg Ser Gln Met
                325                 330                 335
Pro Tyr Lys Val Phe His Tyr Gln Val Lys Ile His Phe Ser Gly Thr
            340                 345                 350
Glu Ser Glu Thr His Thr Asn Gln Ala Phe Glu Ile Ser Leu Tyr Gly
        355                 360                 365
Thr Val Ala Glu Ser Glu Asn Ile Pro Phe Thr Leu Pro Glu Val Ser
    370                 375                 380
Thr Asn Lys Thr Tyr Ser Phe Leu Ile Tyr Thr Glu Val Asp Ile Gly
385                 390                 395                 400
Glu Leu Leu Met Leu Lys Leu Lys Trp Lys Ser Asp Ser Tyr Phe Ser
                405                 410                 415
Trp Ser Asp Trp Trp Ser Ser Pro Gly Phe Ala Ile Gln Lys Ile Arg
            420                 425                 430
Val Lys Ala Gly Glu Thr Gln Lys Lys Val Ile Phe Cys Ser Arg Glu
        435                 440                 445
Lys Val Ser His Leu Gln Lys Gly Lys Ala Pro Ala Val Phe Val Lys
    450                 455                 460
Cys His Asp Lys Ser Leu Asn Lys Lys Ser Gly
465                 470                 475
```

<210> 15
<211> 1466
<212> DNA
<213> Homo sapien

```
<220>
<221> CDS
<222> (115)...(1305)
<223> Nucleotide sequence encoding apolipoprotein A-IV
      (APOA4)

<400> 15
agttcccact gcagcgcagg tgagctctcc tgaggacctc tctgtcagct cccctgattg     60
tagggaggca tccagtgtgg caagaaactc ctccagccca gcaagcagct cagg atg      117
                                                             Met
                                                              1

ttc ctg aag gcc gtg gtc ctg acc ctg gcc ctg gtg gct gtc gcc gga      165
Phe Leu Lys Ala Val Val Leu Thr Leu Ala Leu Val Ala Val Ala Gly
            5                   10                  15

gcc agg gct gag gtc agt gct gac cag gtg gcc aca gtg atg tgg gac      213
Ala Arg Ala Glu Val Ser Ala Asp Gln Val Ala Thr Val Met Trp Asp
        20                  25                  30

tac ttc agc cag ctg agc aac aat gcc aag gag gcc gtg gaa cat ctc      261
Tyr Phe Ser Gln Leu Ser Asn Asn Ala Lys Glu Ala Val Glu His Leu
    35                  40                  45

cag aaa tct gaa ctc acc cag caa ctc aat gcc ctc ttc cag gac aaa      309
Gln Lys Ser Glu Leu Thr Gln Gln Leu Asn Ala Leu Phe Gln Asp Lys
50                  55                  60                  65

ctt gga gaa gtg aac act tac gca ggt gac ctg cag aag aag ctg gtg      357
Leu Gly Glu Val Asn Thr Tyr Ala Gly Asp Leu Gln Lys Lys Leu Val
                70                  75                  80

ccc ttt gcc acc gag ctg cat gaa cgc ctg gcc aag gac tcg gag aaa      405
Pro Phe Ala Thr Glu Leu His Glu Arg Leu Ala Lys Asp Ser Glu Lys
            85                  90                  95

ctg aag gag gag att ggg aag gag ctg gag gag ctg agg gcc cgg ctg      453
Leu Lys Glu Glu Ile Gly Lys Glu Leu Glu Glu Leu Arg Ala Arg Leu
        100                 105                 110

ctg ccc cat gcc aat gag gtg agc cag aag atc ggg gac aac ctg cga      501
Leu Pro His Ala Asn Glu Val Ser Gln Lys Ile Gly Asp Asn Leu Arg
    115                 120                 125

gag ctt cag cag cgc ctg gag ccc tac gcg gac cag ctg cgc acc cag      549
Glu Leu Gln Gln Arg Leu Glu Pro Tyr Ala Asp Gln Leu Arg Thr Gln
130                 135                 140                 145

gtc aac acg cag gcc gag cag ctg cgg cgc cag ctg acc ccc tac gca      597
Val Asn Thr Gln Ala Glu Gln Leu Arg Arg Gln Leu Thr Pro Tyr Ala
                150                 155                 160

cag cgc atg gag aga gtg ctg cgg gag aac gcc gac agc ctg cag gcc      645
Gln Arg Met Glu Arg Val Leu Arg Glu Asn Ala Asp Ser Leu Gln Ala
            165                 170                 175

tcg ctg agg ccc cac gcc gac gag ctc aag gcc aag atc gac cag aac      693
Ser Leu Arg Pro His Ala Asp Glu Leu Lys Ala Lys Ile Asp Gln Asn
```

```
        180                 185                 190

gtg gag gag ctc aag gga cgc ctt acg ccc tac gct gac gaa ttc aaa       741
Val Glu Glu Leu Lys Gly Arg Leu Thr Pro Tyr Ala Asp Glu Phe Lys
    195                 200                 205

gtc aag att gac cag acc gtg gag gag ctg cgc cgc agc ctg gct ccc       789
Val Lys Ile Asp Gln Thr Val Glu Glu Leu Arg Arg Ser Leu Ala Pro
210                 215                 220                 225

tat gct cag gac acg cag gag aag ctc aac cac cag ctt gag ggc ctg       837
Tyr Ala Gln Asp Thr Gln Glu Lys Leu Asn His Gln Leu Glu Gly Leu
                230                 235                 240

acc ttc cag atg aag aag aac gcc gag gag ctc aag gcc agg atc tcg       885
Thr Phe Gln Met Lys Lys Asn Ala Glu Glu Leu Lys Ala Arg Ile Ser
            245                 250                 255

gcc agt gcc gag gag ctg cgg cag agg ctg gcg ccc ttg gcc gag gac       933
Ala Ser Ala Glu Glu Leu Arg Gln Arg Leu Ala Pro Leu Ala Glu Asp
        260                 265                 270

gtg cgt ggc aac ctg agg ggc aac acc gag ggg ctg cag aag tca ctg       981
Val Arg Gly Asn Leu Arg Gly Asn Thr Glu Gly Leu Gln Lys Ser Leu
    275                 280                 285

gca gag ctg ggt ggg cac ctg gac cag cag gtg gag gag ttc cga cgc      1029
Ala Glu Leu Gly Gly His Leu Asp Gln Gln Val Glu Glu Phe Arg Arg
290                 295                 300                 305

cgg gtg gag ccc tac ggg gaa aac ttc aac aaa gcc ctg gtg cag cag      1077
Arg Val Glu Pro Tyr Gly Glu Asn Phe Asn Lys Ala Leu Val Gln Gln
                310                 315                 320

atg gaa cag ctc agg acg aaa ctg ggc ccc cat gcg ggg gac gtg gaa      1125
Met Glu Gln Leu Arg Thr Lys Leu Gly Pro His Ala Gly Asp Val Glu
            325                 330                 335

ggc cac ttg agc ttc ctg gag aag gac ctg agg gac aag gtc aac tcc      1173
Gly His Leu Ser Phe Leu Glu Lys Asp Leu Arg Asp Lys Val Asn Ser
        340                 345                 350

ttc ttc agc acc ttc aag gag aaa gag agc cag gac aag act ctc tcc      1221
Phe Phe Ser Thr Phe Lys Glu Lys Glu Ser Gln Asp Lys Thr Leu Ser
    355                 360                 365

ctc cct gag ctg gag caa cag cag gaa cag cat cag gag cag cag cag      1269
Leu Pro Glu Leu Glu Gln Gln Gln Glu Gln His Gln Glu Gln Gln Gln
370                 375                 380                 385

gag cag gtg cag atg ctg gcc cct ttg gag agc tga gctgcccctg           1315
Glu Gln Val Gln Met Leu Ala Pro Leu Glu Ser  *
                390                 395

gtgcactggc cccaccctcg tggacacctg ccctgccctg ccacctgtct gtctgtccca    1375
aagaagttct ggtatgaact tgaggacaca tgtccagtgg gaggtgagac cacctctcaa    1435
tattcaataa agctgctgag aatctagcct c                                   1466
```

<210> 16
<211> 396
<212> PRT
<213> Homo sapien

<400> 16
Met Phe Leu Lys Ala Val Val Leu Thr Leu Ala Leu Val Ala Val Ala
1 5 10 15
Gly Ala Arg Ala Glu Val Ser Ala Asp Gln Val Ala Thr Val Met Trp
20 25 30
Asp Tyr Phe Ser Gln Leu Ser Asn Asn Ala Lys Glu Ala Val Glu His
35 40 45
Leu Gln Lys Ser Glu Leu Thr Gln Gln Leu Asn Ala Leu Phe Gln Asp
50 55 60
Lys Leu Gly Glu Val Asn Thr Tyr Ala Gly Asp Leu Gln Lys Lys Leu
65 70 75 80
Val Pro Phe Ala Thr Glu Leu His Glu Arg Leu Ala Lys Asp Ser Glu
85 90 95
Lys Leu Lys Glu Glu Ile Gly Lys Glu Leu Glu Glu Leu Arg Ala Arg
100 105 110
Leu Leu Pro His Ala Asn Glu Val Ser Gln Lys Ile Gly Asp Asn Leu
115 120 125
Arg Glu Leu Gln Gln Arg Leu Glu Pro Tyr Ala Asp Gln Leu Arg Thr
130 135 140
Gln Val Asn Thr Gln Ala Glu Gln Leu Arg Arg Gln Leu Thr Pro Tyr
145 150 155 160
Ala Gln Arg Met Glu Arg Val Leu Arg Glu Asn Ala Asp Ser Leu Gln
165 170 175
Ala Ser Leu Arg Pro His Ala Asp Glu Leu Lys Ala Lys Ile Asp Gln
180 185 190
Asn Val Glu Glu Leu Lys Gly Arg Leu Thr Pro Tyr Ala Asp Glu Phe
195 200 205
Lys Val Lys Ile Asp Gln Thr Val Glu Glu Leu Arg Arg Ser Leu Ala
210 215 220
Pro Tyr Ala Gln Asp Thr Gln Glu Lys Leu Asn His Gln Leu Glu Gly
225 230 235 240
Leu Thr Phe Gln Met Lys Lys Asn Ala Glu Glu Leu Lys Ala Arg Ile
245 250 255
Ser Ala Ser Ala Glu Glu Leu Arg Gln Arg Leu Ala Pro Leu Ala Glu
260 265 270
Asp Val Arg Gly Asn Leu Arg Gly Asn Thr Glu Gly Leu Gln Lys Ser
275 280 285
Leu Ala Glu Leu Gly Gly His Leu Asp Gln Gln Val Glu Glu Phe Arg
290 295 300
Arg Arg Val Glu Pro Tyr Gly Glu Asn Phe Asn Lys Ala Leu Val Gln
305 310 315 320
Gln Met Glu Gln Leu Arg Thr Lys Leu Gly Pro His Ala Gly Asp Val
325 330 335
Glu Gly His Leu Ser Phe Leu Glu Lys Asp Leu Arg Asp Lys Val Asn
340 345 350
Ser Phe Phe Ser Thr Phe Lys Glu Lys Glu Ser Gln Asp Lys Thr Leu
355 360 365
Ser Leu Pro Glu Leu Glu Gln Gln Gln Glu Gln His Gln Glu Gln Gln
370 375 380
Gln Glu Gln Val Gln Met Leu Ala Pro Leu Glu Ser
385 390 395

<210> 17

<211> 1156
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (61)...(1014)
<223> Nucleotide Sequence encoding apolipoprotein E
(APOE)

<400> 17
```
cgcagcggag gtgaaggacg tccttcccca ggagccgact ggccaatcac aggcaggaag      60
atg aag gtt ctg tgg gct gcg ttg ctg gtc aca ttc ctg gca gga tgc       108
Met Lys Val Leu Trp Ala Ala Leu Leu Val Thr Phe Leu Ala Gly Cys
 1               5                  10                  15

cag gcc aag gtg gag caa gcg gtg gag aca gag ccg gag ccc gag ctg       156
Gln Ala Lys Val Glu Gln Ala Val Glu Thr Glu Pro Glu Pro Glu Leu
            20                  25                  30

cgc cag cag acc gag tgg cag agc ggc cag cgc tgg gaa ctg gca ctg       204
Arg Gln Gln Thr Glu Trp Gln Ser Gly Gln Arg Trp Glu Leu Ala Leu
        35                  40                  45

ggt cgc ttt tgg gat tac ctg cgc tgg gtg cag aca ctg tct gag cag       252
Gly Arg Phe Trp Asp Tyr Leu Arg Trp Val Gln Thr Leu Ser Glu Gln
    50                  55                  60

gtg cag gag gag ctg ctc agc tcc cag gtc acc cag gaa ctg agg gcg       300
Val Gln Glu Glu Leu Leu Ser Ser Gln Val Thr Gln Glu Leu Arg Ala
 65                  70                  75                  80

ctg atg gac gag acc atg aag gag ttg aag gcc tac aaa tcg gaa ctg       348
Leu Met Asp Glu Thr Met Lys Glu Leu Lys Ala Tyr Lys Ser Glu Leu
                85                  90                  95

gag gaa caa ctg acc ccg gtg gcg gag gag acg cgg gca cgg ctg tcc       396
Glu Glu Gln Leu Thr Pro Val Ala Glu Glu Thr Arg Ala Arg Leu Ser
            100                 105                 110

aag gag ctg cag gcg gcg cag gcc cgg ctg ggc gcg gac atg gag gac       444
Lys Glu Leu Gln Ala Ala Gln Ala Arg Leu Gly Ala Asp Met Glu Asp
        115                 120                 125

gtg tgc ggc cgc ctg gtg cag tac cgc ggc gag gtg cag gcc atg ctc       492
Val Cys Gly Arg Leu Val Gln Tyr Arg Gly Glu Val Gln Ala Met Leu
    130                 135                 140

ggc cag agc acc gag gag ctg cgg gtg cgc ctc gcc tcc cac ctg cgc       540
Gly Gln Ser Thr Glu Glu Leu Arg Val Arg Leu Ala Ser His Leu Arg
145                 150                 155                 160

aag ctg cgt aag cgg ctc ctc cgc gat gcc gat gac ctg cag aag cgc       588
Lys Leu Arg Lys Arg Leu Leu Arg Asp Ala Asp Asp Leu Gln Lys Arg
                165                 170                 175

ctg gca gtg tac cag gcc ggg gcc cgc gag ggc gcc gag cgc ggc ctc       636
Leu Ala Val Tyr Gln Ala Gly Ala Arg Glu Gly Ala Glu Arg Gly Leu
```

```
            180                 185                 190

agc gcc atc cgc gag cgc ctg ggg ccc ctg gtg gaa cag ggc cgc gtg        684
Ser Ala Ile Arg Glu Arg Leu Gly Pro Leu Val Glu Gln Gly Arg Val
        195                 200                 205

cgg gcc gcc act gtg ggc tcc ctg gcc ggc cag ccg cta cag gag cgg        732
Arg Ala Ala Thr Val Gly Ser Leu Ala Gly Gln Pro Leu Gln Glu Arg
    210                 215                 220

gcc cag gcc tgg ggc gag cgg ctg cgc gcg cgg atg gag gag atg ggc        780
Ala Gln Ala Trp Gly Glu Arg Leu Arg Ala Arg Met Glu Glu Met Gly
225                 230                 235                 240

agc cgg acc cgc gac cgc ctg gac gag gtg aag gag cag gtg gcg gag        828
Ser Arg Thr Arg Asp Arg Leu Asp Glu Val Lys Glu Gln Val Ala Glu
                245                 250                 255

gtg cgc gcc aag ctg gag gag cag gcc cag cag ata cgc ctg cag gcc        876
Val Arg Ala Lys Leu Glu Glu Gln Ala Gln Gln Ile Arg Leu Gln Ala
            260                 265                 270

gag gcc ttc cag gcc cgc ctc aag agc tgg ttc gag ccc ctg gtg gaa        924
Glu Ala Phe Gln Ala Arg Leu Lys Ser Trp Phe Glu Pro Leu Val Glu
        275                 280                 285

gac atg cag cgc cag tgg gcc ggg ctg gtg gag aag gtg cag gct gcc        972
Asp Met Gln Arg Gln Trp Ala Gly Leu Val Glu Lys Val Gln Ala Ala
    290                 295                 300

gtg ggc acc agc gcc gcc cct gtg ccc agc gac aat cac tga               1014
Val Gly Thr Ser Ala Ala Pro Val Pro Ser Asp Asn His  *
305                 310                 315

acgccgaagc ctgcagccat gcgaccccac gccaccccgt gcctcctgcc tccgcgcagc    1074
ctgcagcggg agaccctgtc cccgccccag ccgtcctcct ggggtggacc ctagtttaat    1134
aaagattcac caagtttcac gc                                             1156

<210> 18
<211> 317
<212> PRT
<213> Homo sapien

<400> 18
Met Lys Val Leu Trp Ala Ala Leu Leu Val Thr Phe Leu Ala Gly Cys
 1               5                  10                  15
Gln Ala Lys Val Glu Gln Ala Val Glu Thr Glu Pro Glu Pro Glu Leu
            20                  25                  30
Arg Gln Gln Thr Glu Trp Gln Ser Gly Gln Arg Trp Glu Leu Ala Leu
        35                  40                  45
Gly Arg Phe Trp Asp Tyr Leu Arg Trp Val Gln Thr Leu Ser Glu Gln
    50                  55                  60
Val Gln Glu Glu Leu Leu Ser Ser Gln Val Thr Gln Glu Leu Arg Ala
65                  70                  75                  80
Leu Met Asp Glu Thr Met Lys Glu Leu Lys Ala Tyr Lys Ser Glu Leu
                85                  90                  95
Glu Glu Gln Leu Thr Pro Val Ala Glu Glu Thr Arg Ala Arg Leu Ser
            100                 105                 110
```

```
Lys Glu Leu Gln Ala Ala Gln Ala Arg Leu Gly Ala Asp Met Glu Asp
        115                 120                 125
Val Cys Gly Arg Leu Val Gln Tyr Arg Gly Glu Val Gln Ala Met Leu
    130                 135                 140
Gly Gln Ser Thr Glu Glu Leu Arg Val Arg Leu Ala Ser His Leu Arg
145                 150                 155                 160
Lys Leu Arg Lys Arg Leu Leu Arg Asp Ala Asp Asp Leu Gln Lys Arg
                165                 170                 175
Leu Ala Val Tyr Gln Ala Gly Ala Arg Glu Gly Ala Glu Arg Gly Leu
            180                 185                 190
Ser Ala Ile Arg Glu Arg Leu Gly Pro Leu Val Glu Gln Gly Arg Val
        195                 200                 205
Arg Ala Ala Thr Val Gly Ser Leu Ala Gly Gln Pro Leu Gln Glu Arg
    210                 215                 220
Ala Gln Ala Trp Gly Glu Arg Leu Arg Ala Arg Met Glu Glu Met Gly
225                 230                 235                 240
Ser Arg Thr Arg Asp Arg Leu Asp Glu Val Lys Glu Gln Val Ala Glu
                245                 250                 255
Val Arg Ala Lys Leu Glu Glu Gln Ala Gln Gln Ile Arg Leu Gln Ala
            260                 265                 270
Glu Ala Phe Gln Ala Arg Leu Lys Ser Trp Phe Glu Pro Leu Val Glu
        275                 280                 285
Asp Met Gln Arg Gln Trp Ala Gly Leu Val Glu Lys Val Gln Ala Ala
    290                 295                 300
Val Gly Thr Ser Ala Ala Pro Val Pro Ser Asp Asn His
305                 310                 315

<210> 19
<211> 1603
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (58)...(1557)
<223> Nucleotide sequence encoding hepatic lipase (LIPC)

<400> 19
ggtctctttg gcttcagaaa ttaccaagaa agcctggacc ccgggtgaaa cggagaa atg      60
                                                               Met
                                                                1

gac aca agt ccc ctg tgt ttc tcc att ctg ttg gtt tta tgc atc ttt      108
Asp Thr Ser Pro Leu Cys Phe Ser Ile Leu Leu Val Leu Cys Ile Phe
            5                  10                  15

atc caa tca agt gcc ctt gga caa agc ctg aaa cca gag cca ttt gga      156
Ile Gln Ser Ser Ala Leu Gly Gln Ser Leu Lys Pro Glu Pro Phe Gly
        20                  25                  30

aga aga gct caa gct gtt gaa aca aac aaa acg ctg cat gag atg aag      204
Arg Arg Ala Gln Ala Val Glu Thr Asn Lys Thr Leu His Glu Met Lys
    35                  40                  45

acc aga ttc ctg ctc ttt gga gaa acc aat cag ggc tgt cag att cga      252
Thr Arg Phe Leu Leu Phe Gly Glu Thr Asn Gln Gly Cys Gln Ile Arg
 50                  55                  60                  65
```

```
atc aat cat ccg gac acg tta cag gag tgc ggc ttc aac tcc tcc ctg        300
Ile Asn His Pro Asp Thr Leu Gln Glu Cys Gly Phe Asn Ser Ser Leu
                 70                  75                  80

cct ctg gtg atg ata atc cac ggg tgg tcg gtg gac ggc gtg cta gaa        348
Pro Leu Val Met Ile Ile His Gly Trp Ser Val Asp Gly Val Leu Glu
             85                  90                  95

aac tgg atc tgg cag atg gtg gcc gcg ctg aag tct cag ccg gcc cag        396
Asn Trp Ile Trp Gln Met Val Ala Ala Leu Lys Ser Gln Pro Ala Gln
        100                 105                 110

cca gtg aac gtg ggg ctg gtg gac tgg atc acc ctg gcc cac gac cac        444
Pro Val Asn Val Gly Leu Val Asp Trp Ile Thr Leu Ala His Asp His
    115                 120                 125

tac acc atc gcc gtc cgc aac acc cgc ctt gtg ggc aag gag gtc gcg        492
Tyr Thr Ile Ala Val Arg Asn Thr Arg Leu Val Gly Lys Glu Val Ala
130                 135                 140                 145

gct ctt ctc cgg tgg ctg gag gaa tct gtt caa ctc tct cga agc cat        540
Ala Leu Leu Arg Trp Leu Glu Glu Ser Val Gln Leu Ser Arg Ser His
                150                 155                 160

gtt cac cta att ggg tac agc ctg ggt gca cac gtg tca gga ttt gcc        588
Val His Leu Ile Gly Tyr Ser Leu Gly Ala His Val Ser Gly Phe Ala
            165                 170                 175

ggc agt tcc atc ggt gga acg cac aag att ggg aga atc aca ggg ctg        636
Gly Ser Ser Ile Gly Gly Thr His Lys Ile Gly Arg Ile Thr Gly Leu
        180                 185                 190

gat gcc gcg gga cct ttg ttt gag gga agt gcc ccc agc aat cgt ctt        684
Asp Ala Ala Gly Pro Leu Phe Glu Gly Ser Ala Pro Ser Asn Arg Leu
    195                 200                 205

tct cca gat gat gcc aat ttt gtg gat gcc att cat acc ttt acg cgg        732
Ser Pro Asp Asp Ala Asn Phe Val Asp Ala Ile His Thr Phe Thr Arg
210                 215                 220                 225

gag cac atg ggc ctg agc gtg ggc atc aaa cag ccc ata gga cac tat        780
Glu His Met Gly Leu Ser Val Gly Ile Lys Gln Pro Ile Gly His Tyr
                230                 235                 240

gac ttc tat ccc aac ggg ggc tcc ttc cag cct ggc tgc cac ttc cta        828
Asp Phe Tyr Pro Asn Gly Gly Ser Phe Gln Pro Gly Cys His Phe Leu
            245                 250                 255

gag ctc tac aga cat att gcc cag cac ggc ttc aat gcc atc acc cag        876
Glu Leu Tyr Arg His Ile Ala Gln His Gly Phe Asn Ala Ile Thr Gln
        260                 265                 270

acc ata aaa tgc tcc cac gag cga tcg gtg cac ctt ttc atc gac tcc        924
Thr Ile Lys Cys Ser His Glu Arg Ser Val His Leu Phe Ile Asp Ser
    275                 280                 285

ttg ctg cac gcc ggc acg cag agc atg gcc tac ccg tgt ggt gac atg        972
Leu Leu His Ala Gly Thr Gln Ser Met Ala Tyr Pro Cys Gly Asp Met
```

```
290                 295                 300                 305

aac agc ttc agc cag ggc ctg tgc ctg agc tgc aag aag ggc cgc tgc      1020
Asn Ser Phe Ser Gln Gly Leu Cys Leu Ser Cys Lys Lys Gly Arg Cys
                310                 315                 320

aac acg ctg ggc tac cac gtc cgc cag gag ccg cgg agc aag agc aag      1068
Asn Thr Leu Gly Tyr His Val Arg Gln Glu Pro Arg Ser Lys Ser Lys
            325                 330                 335

agg ctc ttc ctc gta acg cga gcc cag tcc ccc ttc aaa gtt tat cat      1116
Arg Leu Phe Leu Val Thr Arg Ala Gln Ser Pro Phe Lys Val Tyr His
        340                 345                 350

tac cag tta aag atc cag ttc atc aac caa act gag acg cca ata caa      1164
Tyr Gln Leu Lys Ile Gln Phe Ile Asn Gln Thr Glu Thr Pro Ile Gln
    355                 360                 365

aca act ttt acc atg tca cta ctc gga aca aaa gag aaa atg cag aaa      1212
Thr Thr Phe Thr Met Ser Leu Leu Gly Thr Lys Glu Lys Met Gln Lys
370                 375                 380                 385

att ccc atc act ctg ggc aaa gga att gct agt aat aaa acg tat tcc      1260
Ile Pro Ile Thr Leu Gly Lys Gly Ile Ala Ser Asn Lys Thr Tyr Ser
                390                 395                 400

ttt ctt atc acg ctg gat gtg gat atc ggc gag ctg atc atg atc aag      1308
Phe Leu Ile Thr Leu Asp Val Asp Ile Gly Glu Leu Ile Met Ile Lys
            405                 410                 415

ttc aag tgg gaa aac agt gca gtg tgg gcc aat gtc tgg gac acg gtc      1356
Phe Lys Trp Glu Asn Ser Ala Val Trp Ala Asn Val Trp Asp Thr Val
        420                 425                 430

cag acc atc atc cca tgg agc aca ggg ccg cgc cac tca ggc ctc gtt      1404
Gln Thr Ile Ile Pro Trp Ser Thr Gly Pro Arg His Ser Gly Leu Val
    435                 440                 445

ctg aag acg atc aga gtc aaa gca gga gaa acc cag caa aga atg aca      1452
Leu Lys Thr Ile Arg Val Lys Ala Gly Glu Thr Gln Gln Arg Met Thr
450                 455                 460                 465

ttt tgt tca gaa aac aca gat gac cta cta ctt cgc cca acc cag gaa      1500
Phe Cys Ser Glu Asn Thr Asp Asp Leu Leu Leu Arg Pro Thr Gln Glu
                470                 475                 480

aaa atc ttc gtg aaa tgt gaa ata aag tct aaa aca tca aag cga aag      1548
Lys Ile Phe Val Lys Cys Glu Ile Lys Ser Lys Thr Ser Lys Arg Lys
            485                 490                 495

atc aga tga gatttaatga agacccagtg taaagaataa atgaatctta              1597
Ile Arg  *

ctcctt                                                                1603

<210> 20
<211> 499
```

<212> PRT
<213> Homo sapien

<400> 20
Met Asp Thr Ser Pro Leu Cys Phe Ser Ile Leu Leu Val Leu Cys Ile
1 5 10 15
Phe Ile Gln Ser Ser Ala Leu Gly Gln Ser Leu Lys Pro Glu Pro Phe
20 25 30
Gly Arg Arg Ala Gln Ala Val Glu Thr Asn Lys Thr Leu His Glu Met
35 40 45
Lys Thr Arg Phe Leu Leu Phe Gly Glu Thr Asn Gln Gly Cys Gln Ile
50 55 60
Arg Ile Asn His Pro Asp Thr Leu Gln Glu Cys Gly Phe Asn Ser Ser
65 70 75 80
Leu Pro Leu Val Met Ile Ile His Gly Trp Ser Val Asp Gly Val Leu
85 90 95
Glu Asn Trp Ile Trp Gln Met Val Ala Ala Leu Lys Ser Gln Pro Ala
100 105 110
Gln Pro Val Asn Val Gly Leu Val Asp Trp Ile Thr Leu Ala His Asp
115 120 125
His Tyr Thr Ile Ala Val Arg Asn Thr Arg Leu Val Gly Lys Glu Val
130 135 140
Ala Ala Leu Leu Arg Trp Leu Glu Glu Ser Val Gln Leu Ser Arg Ser
145 150 155 160
His Val His Leu Ile Gly Tyr Ser Leu Gly Ala His Val Ser Gly Phe
165 170 175
Ala Gly Ser Ser Ile Gly Gly Thr His Lys Ile Gly Arg Ile Thr Gly
180 185 190
Leu Asp Ala Ala Gly Pro Leu Phe Glu Gly Ser Ala Pro Ser Asn Arg
195 200 205
Leu Ser Pro Asp Asp Ala Asn Phe Val Asp Ala Ile His Thr Phe Thr
210 215 220
Arg Glu His Met Gly Leu Ser Val Gly Ile Lys Gln Pro Ile Gly His
225 230 235 240
Tyr Asp Phe Tyr Pro Asn Gly Gly Ser Phe Gln Pro Gly Cys His Phe
245 250 255
Leu Glu Leu Tyr Arg His Ile Ala Gln His Gly Phe Asn Ala Ile Thr
260 265 270
Gln Thr Ile Lys Cys Ser His Glu Arg Ser Val His Leu Phe Ile Asp
275 280 285
Ser Leu Leu His Ala Gly Thr Gln Ser Met Ala Tyr Pro Cys Gly Asp
290 295 300
Met Asn Ser Phe Ser Gln Gly Leu Cys Leu Ser Cys Lys Lys Gly Arg
305 310 315 320
Cys Asn Thr Leu Gly Tyr His Val Arg Gln Glu Pro Arg Ser Lys Ser
325 330 335
Lys Arg Leu Phe Leu Val Thr Arg Ala Gln Ser Pro Phe Lys Val Tyr
340 345 350
His Tyr Gln Leu Lys Ile Gln Phe Ile Asn Gln Thr Glu Thr Pro Ile
355 360 365
Gln Thr Thr Phe Thr Met Ser Leu Leu Gly Thr Lys Glu Lys Met Gln
370 375 380
Lys Ile Pro Ile Thr Leu Gly Lys Gly Ile Ala Ser Asn Lys Thr Tyr
385 390 395 400
Ser Phe Leu Ile Thr Leu Asp Val Asp Ile Gly Glu Leu Ile Met Ile
405 410 415
Lys Phe Lys Trp Glu Asn Ser Ala Val Trp Ala Asn Val Trp Asp Thr
420 425 430

```
Val Gln Thr Ile Ile Pro Trp Ser Thr Gly Pro Arg His Ser Gly Leu
        435                 440                 445
Val Leu Lys Thr Ile Arg Val Lys Ala Gly Glu Thr Gln Gln Arg Met
    450                 455                 460
Thr Phe Cys Ser Glu Asn Thr Asp Asp Leu Leu Leu Arg Pro Thr Gln
465                 470                 475                 480
Glu Lys Ile Phe Val Lys Cys Glu Ile Lys Ser Lys Thr Ser Lys Arg
                485                 490                 495
Lys Ile Arg


<210> 21
<211> 1346
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (10)...(1077)
<223> Nucleotide sequence encoding paraoxonase 1 (PON1)

<400> 21
cccccgacc atg gcg aag ctg att gcg ctc acc ctc ttg ggg atg gga ctg      51
          Met Ala Lys Leu Ile Ala Leu Thr Leu Leu Gly Met Gly Leu
           1               5                  10

gca ctc ttc agg aac cac cag tct tct tac caa aca cga ctt aat gct        99
Ala Leu Phe Arg Asn His Gln Ser Ser Tyr Gln Thr Arg Leu Asn Ala
 15                  20                  25                  30

ctc cga gag gta caa ccc gta gaa ctt cct aac tgt aat tta gtt aaa       147
Leu Arg Glu Val Gln Pro Val Glu Leu Pro Asn Cys Asn Leu Val Lys
                 35                  40                  45

gga atc gaa act ggc tct gaa gac atg gag ata ctg cct aat gga ctg       195
Gly Ile Glu Thr Gly Ser Glu Asp Met Glu Ile Leu Pro Asn Gly Leu
             50                  55                  60

gct ttc att agc tct gga tta aag tat cct gga ata aag agc ttc aac       243
Ala Phe Ile Ser Ser Gly Leu Lys Tyr Pro Gly Ile Lys Ser Phe Asn
         65                  70                  75

ccc aac agt cct gga aaa ata ctt ctg atg gac ctg aat gaa gaa gat       291
Pro Asn Ser Pro Gly Lys Ile Leu Leu Met Asp Leu Asn Glu Glu Asp
     80                  85                  90

cca aca gtg ttg gaa ttg ggg atc act gga agt aaa ttt gat gta tct       339
Pro Thr Val Leu Glu Leu Gly Ile Thr Gly Ser Lys Phe Asp Val Ser
 95                 100                 105                 110

tca ttt aac cct cat ggg att agc aca ttc aca gat gaa gat aat gcc       387
Ser Phe Asn Pro His Gly Ile Ser Thr Phe Thr Asp Glu Asp Asn Ala
                115                 120                 125

atg tac ctc ctg gtg gtg aac cat cca gat gcc aag tcc aca gtg gag       435
Met Tyr Leu Leu Val Val Asn His Pro Asp Ala Lys Ser Thr Val Glu
            130                 135                 140
```

```
ttg ttt aaa ttt caa gaa gaa gaa aaa tcg ctt ttg cat cta aaa acc      483
Leu Phe Lys Phe Gln Glu Glu Glu Lys Ser Leu Leu His Leu Lys Thr
        145                 150                 155

atc aga cat aaa ctt ctg cct aat ttg aat gat att gtt gct gtg gga      531
Ile Arg His Lys Leu Leu Pro Asn Leu Asn Asp Ile Val Ala Val Gly
    160                 165                 170

cct gag cac ttt tat ggc aca aat gat cac tat ttt ctt gac ccc tac      579
Pro Glu His Phe Tyr Gly Thr Asn Asp His Tyr Phe Leu Asp Pro Tyr
175                 180                 185                 190

tta caa tcc tgg gag atg tat ttg ggt tta gcg tgg tcg tat gtt gtc      627
Leu Gln Ser Trp Glu Met Tyr Leu Gly Leu Ala Trp Ser Tyr Val Val
                195                 200                 205

tac tat agt cca agt gaa gtt cga gtg gtg gca gaa gga ttt gat ttt      675
Tyr Tyr Ser Pro Ser Glu Val Arg Val Val Ala Glu Gly Phe Asp Phe
            210                 215                 220

gct aat gga atc aac att tca ccc gat ggc aag tat gtc tat ata gct      723
Ala Asn Gly Ile Asn Ile Ser Pro Asp Gly Lys Tyr Val Tyr Ile Ala
        225                 230                 235

gag ttg ctg gct cat aag att cat gtg tat gaa aag cat gct aat tgg      771
Glu Leu Leu Ala His Lys Ile His Val Tyr Glu Lys His Ala Asn Trp
    240                 245                 250

act tta act cca ttg aag tcc ctt gac ttt aat acc ctc gtg gat aac      819
Thr Leu Thr Pro Leu Lys Ser Leu Asp Phe Asn Thr Leu Val Asp Asn
255                 260                 265                 270

ata tct gtg gat cct gag aca gga gac ctt tgg gtt gga tgc cat ccc      867
Ile Ser Val Asp Pro Glu Thr Gly Asp Leu Trp Val Gly Cys His Pro
                275                 280                 285

aat ggc atg aaa atc ttc ttc tat gac tca gag aat cct cct gca tca      915
Asn Gly Met Lys Ile Phe Phe Tyr Asp Ser Glu Asn Pro Pro Ala Ser
            290                 295                 300

gag gtg ctt cga atc cag aac att cta aca gaa gaa cct aaa gtg aca      963
Glu Val Leu Arg Ile Gln Asn Ile Leu Thr Glu Glu Pro Lys Val Thr
        305                 310                 315

cag gtt tat gca gaa aat ggc aca gtg ttg caa ggc agt aca gtt gcc     1011
Gln Val Tyr Ala Glu Asn Gly Thr Val Leu Gln Gly Ser Thr Val Ala
    320                 325                 330

tct gtg tac aaa ggg aaa ctg ctg att ggc aca gtg ttt cac aaa gct     1059
Ser Val Tyr Lys Gly Lys Leu Leu Ile Gly Thr Val Phe His Lys Ala
335                 340                 345                 350

ctt tac tgt gag ctc taa cagaccgatt tgcacccatg ccatagaaac            1107
Leu Tyr Cys Glu Leu  *
                355

tgaggccatt atttcaaccg cttgccatat tccgaggacc cagtgttctt agctgaacaa   1167
tgaatgctga ccctaaatgt ggacatcatg aagcatcaaa gcactgttta actgggagtg   1227
```

```
atatgatgtg tagggctttt ttttgagaat acactatcaa atcagtcttg gaatacttga    1287
aaacctcatt taccataaaa atccttctca ctaaaatgga taaatcagtt aaaaaaaaa     1346

<210> 22
<211> 355
<212> PRT
<213> Homo sapien

<400> 22
Met Ala Lys Leu Ile Ala Leu Thr Leu Leu Gly Met Gly Leu Ala Leu
 1               5                  10                  15
Phe Arg Asn His Gln Ser Ser Tyr Gln Thr Arg Leu Asn Ala Leu Arg
            20                  25                  30
Glu Val Gln Pro Val Glu Leu Pro Asn Cys Asn Leu Val Lys Gly Ile
        35                  40                  45
Glu Thr Gly Ser Glu Asp Met Glu Ile Leu Pro Asn Gly Leu Ala Phe
    50                  55                  60
Ile Ser Ser Gly Leu Lys Tyr Pro Gly Ile Lys Ser Phe Asn Pro Asn
65                  70                  75                  80
Ser Pro Gly Lys Ile Leu Leu Met Asp Leu Asn Glu Glu Asp Pro Thr
                85                  90                  95
Val Leu Glu Leu Gly Ile Thr Gly Ser Lys Phe Asp Val Ser Ser Phe
            100                 105                 110
Asn Pro His Gly Ile Ser Thr Phe Thr Asp Glu Asp Asn Ala Met Tyr
        115                 120                 125
Leu Leu Val Val Asn His Pro Asp Ala Lys Ser Thr Val Glu Leu Phe
    130                 135                 140
Lys Phe Gln Glu Glu Glu Lys Ser Leu Leu His Leu Lys Thr Ile Arg
145                 150                 155                 160
His Lys Leu Leu Pro Asn Leu Asn Asp Ile Val Ala Val Gly Pro Glu
                165                 170                 175
His Phe Tyr Gly Thr Asn Asp His Tyr Phe Leu Asp Pro Tyr Leu Gln
            180                 185                 190
Ser Trp Glu Met Tyr Leu Gly Leu Ala Trp Ser Tyr Val Val Tyr Tyr
        195                 200                 205
Ser Pro Ser Glu Val Arg Val Val Ala Glu Gly Phe Asp Phe Ala Asn
    210                 215                 220
Gly Ile Asn Ile Ser Pro Asp Gly Lys Tyr Val Tyr Ile Ala Glu Leu
225                 230                 235                 240
Leu Ala His Lys Ile His Val Tyr Glu Lys His Ala Asn Trp Thr Leu
                245                 250                 255
Thr Pro Leu Lys Ser Leu Asp Phe Asn Thr Leu Val Asp Asn Ile Ser
            260                 265                 270
Val Asp Pro Glu Thr Gly Asp Leu Trp Val Gly Cys His Pro Asn Gly
        275                 280                 285
Met Lys Ile Phe Phe Tyr Asp Ser Glu Asn Pro Pro Ala Ser Glu Val
    290                 295                 300
Leu Arg Ile Gln Asn Ile Leu Thr Glu Glu Pro Lys Val Thr Gln Val
305                 310                 315                 320
Tyr Ala Glu Asn Gly Thr Val Leu Gln Gly Ser Thr Val Ala Ser Val
                325                 330                 335
Tyr Lys Gly Lys Leu Leu Ile Gly Thr Val Phe His Lys Ala Leu Tyr
            340                 345                 350
Cys Glu Leu
        355

<210> 23
<211> 1570
```

```
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (1)...(1097)
<223> Nucleotide sequence encoding paraoxonase 2 (PON2)

<400> 23
cgg agc gag gca gcg cgc ccg gct ccc gcg cca tgg ggc ggc tgg tgg       48
Arg Ser Glu Ala Ala Arg Pro Ala Pro Ala Pro Trp Gly Gly Trp Trp
 1               5                  10                  15

ctg tgg gct tgc tgg gga tcg cgc tgg cgc tcc tgg gcg aga ggc ttc       96
Leu Trp Ala Cys Trp Gly Ser Arg Trp Arg Ser Trp Ala Arg Gly Phe
            20                  25                  30

tgg cac tca gaa atc gac tta aag cct cca gag aag tag aat ctg tag      144
Trp His Ser Glu Ile Asp Leu Lys Pro Pro Glu Lys  *  Asn Leu  *
        35                  40                  45

acc ttc cac act gcc acc tga tta aag gaa ttg aag ctg gct ctg aag      192
Thr Phe His Thr Ala Thr  *  Leu Lys Glu Leu Lys Leu Ala Leu Lys
            50                  55                  60

ata ttg aca tac ttc cca atg gtc tgg ctt ttt tta gtg tgg gtc taa      240
Ile Leu Thr Tyr Phe Pro Met Val Trp Leu Phe Leu Val Trp Val  *
            65                  70                  75

aat tcc cag gac tcc aca gct ttg cac cag ata agc ctg gag gaa tac      288
Asn Ser Gln Asp Ser Thr Ala Leu His Gln Ile Ser Leu Glu Glu Tyr
            80                  85                  90

taa tga tgg atc taa aag aag aaa aac caa ggg cac ggg aat taa gaa      336
 *   *  Trp Ile  *  Lys Lys Lys Asn Gln Gly His Gly Asn  *  Glu
                    95                 100

tca gtc gtg ggt ttg att tgg cct cat tca atc cac atg gca tca gca      384
Ser Val Val Gly Leu Ile Trp Pro His Ser Ile His Met Ala Ser Ala
105                 110                 115                 120

ctt tca tag aca acg atg aca cag ttt atc tct ttg ttg taa acc acc      432
Leu Ser  *  Thr Thr Met Thr Gln Phe Ile Ser Leu Leu  *  Thr Thr
                    125                 130

cag aat tca aga ata cag tgg aaa ttt tta aat ttg aag aag cag aaa      480
Gln Asn Ser Arg Ile Gln Trp Lys Phe Leu Asn Leu Lys Lys Gln Lys
135                 140                 145                 150

att ctc tgt tgc atc tga aaa cag tca aac atg agc ttc ttc caa gtg      528
Ile Leu Cys Cys Ile  *  Lys Gln Ser Asn Met Ser Phe Phe Gln Val
                155                 160                 165

tga atg aca tca cag ctg ttg gac cgg cac att tct atg cca caa atg      576
 *  Met Thr Ser Gln Leu Leu Asp Arg His Ile Ser Met Pro Gln Met
                    170                 175                 180

acc act act tct ctg atc ctt tct taa agt att tag aaa cat act tga      624
```

```
Thr Thr Thr Ser Leu Ile Leu Ser  *  Ser Ile  *  Lys His Thr  *
                185                     190

act tac act ggg caa atg ttg ttt act aca gtc caa atg aag tta aag      672
Thr Tyr Thr Gly Gln Met Leu Phe Thr Thr Val Gln Met Lys Leu Lys
    195                 200                 205

tgg tag cag aag gat ttg att cag caa atg gga tca ata ttt cac ctg      720
Trp  *  Gln Lys Asp Leu Ile Gln Gln Met Gly Ser Ile Phe His Leu
210                     215                 220

atg ata agt ata tct atg ttg ctg aca tat tgg ctc atg aaa ttc atg      768
Met Ile Ser Ile Ser Met Leu Leu Thr Tyr Trp Leu Met Lys Phe Met
225                 230                 235                 240

ttt tgg aaa aac aca cta ata tga att taa ctc agt tga agg tac ttg      816
Phe Trp Lys Asn Thr Leu Ile  *  Ile  *  Leu Ser  *  Arg Tyr Leu
                245                     250

agc tgg ata cac tgg tgg ata att tat cta ttg atc ctt cct cgg ggg      864
Ser Trp Ile His Trp Trp Ile Ile Tyr Leu Leu Ile Leu Pro Arg Gly
    255                 260                 265

aca tct ggg tag gct gtc atc cta atg gcc aga agc tct tcg tgt atg      912
Thr Ser Gly  *  Ala Val Ile Leu Met Ala Arg Ser Ser Ser Cys Met
270                     275                 280

acc cga aca atc ctc cct cgt cag agg ttc tcc gca tcc aga aca ttc      960
Thr Arg Thr Ile Leu Pro Arg Gln Arg Phe Ser Ala Ser Arg Thr Phe
285                 290                 295                 300

tat ctg aga agc cta cag tga cta cag ttt atg cca aca atg ggt ctg     1008
Tyr Leu Arg Ser Leu Gln  *  Leu Gln Phe Met Pro Thr Met Gly Leu
                305                 310                     315

ttc tcc aag gaa gtt ctg tag cct cag tgt atg atg gga agc tgc tca     1056
Phe Ser Lys Glu Val Leu  *  Pro Gln Cys Met Met Gly Ser Cys Ser
                320                 325                     330

tag gca ctt tat acc aca gag cct tgt att gtg aac tct aa attgtacttt   1107
 *  Ala Leu Tyr Thr Thr Glu Pro Cys Ile Val Asn Ser
                    335                 340

tggcatgaaa gtgcgataac ttaacaatta attttctatg aattgctaat tctgagggaa   1167
tttaaccagc aacattgacc cagaaatgta tggcatgtgt agttaatttt attccagtaa   1227
ggaacggccc ttttagttct tagagcactt ttaacaaaaa aggaaaatga acaggttctt   1287
taaaatgcca agcaagggac agaaaagaaa gctgctttcg aataaagtga atacattttg   1347
cacaaagtaa gcctcacctt tgccttccaa ctgccagaac atggattcca ctgaaataga   1407
gtgaattata tttccttaaa atgtgagtga cctcacttct ggcactgtga ctactatggc   1467
tgtttagaac tactgataac gtattttgat gttttgtact tacatctttg tttaccatta   1527
aaaagttgga gttatattaa agactaacta aaatcccagt ttt                     1570

<210> 24
<211> 342
<212> PRT
<213> Homo sapien

<400> 24
```

```
Arg Ser Glu Ala Ala Arg Pro Ala Pro Ala Pro Trp Gly Gly Trp Trp
 1               5                  10                  15
Leu Trp Ala Cys Trp Gly Ser Arg Trp Arg Ser Trp Ala Arg Gly Phe
            20                  25                  30
Trp His Ser Glu Ile Asp Leu Lys Pro Pro Glu Lys Asn Leu Thr Phe
        35                  40                  45
His Thr Ala Thr Leu Lys Glu Leu Lys Leu Ala Leu Lys Ile Leu Thr
    50                  55                  60
Tyr Phe Pro Met Val Trp Leu Phe Leu Val Trp Val Asn Ser Gln Asp
65                  70                  75                  80
Ser Thr Ala Leu His Gln Ile Ser Leu Glu Glu Tyr Trp Ile Lys Lys
                85                  90                  95
Lys Asn Gln Gly His Gly Asn Glu Ser Val Val Gly Leu Ile Trp Pro
            100                 105                 110
His Ser Ile His Met Ala Ser Ala Leu Ser Thr Thr Met Thr Gln Phe
        115                 120                 125
Ile Ser Leu Leu Thr Thr Gln Asn Ser Arg Ile Gln Trp Lys Phe Leu
    130                 135                 140
Asn Leu Lys Lys Gln Lys Ile Leu Cys Cys Ile Lys Gln Ser Asn Met
145                 150                 155                 160
Ser Phe Phe Gln Val Met Thr Ser Gln Leu Leu Asp Arg His Ile Ser
                165                 170                 175
Met Pro Gln Met Thr Thr Thr Ser Leu Ile Leu Ser Ser Ile Lys His
            180                 185                 190
Thr Thr Tyr Thr Gly Gln Met Leu Phe Thr Thr Val Gln Met Lys Leu
        195                 200                 205
Lys Trp Gln Lys Asp Leu Ile Gln Gln Met Gly Ser Ile Phe His Leu
    210                 215                 220
Met Ile Ser Ile Ser Met Leu Leu Thr Tyr Trp Leu Met Lys Phe Met
225                 230                 235                 240
Phe Trp Lys Asn Thr Leu Ile Ile Leu Ser Arg Tyr Leu Ser Trp Ile
                245                 250                 255
His Trp Trp Ile Ile Tyr Leu Leu Ile Leu Pro Arg Gly Thr Ser Gly
            260                 265                 270
Ala Val Ile Leu Met Ala Arg Ser Ser Ser Cys Met Thr Arg Thr Ile
        275                 280                 285
Leu Pro Arg Gln Arg Phe Ser Ala Ser Arg Thr Phe Tyr Leu Arg Ser
    290                 295                 300
Leu Gln Leu Gln Phe Met Pro Thr Met Gly Leu Phe Ser Lys Glu Val
305                 310                 315                 320
Leu Pro Gln Cys Met Met Gly Ser Cys Ser Ala Leu Tyr Thr Thr Glu
                325                 330                 335
Pro Cys Ile Val Asn Ser
            340
```

<210> 25
<211> 533
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (47)...(346)
<223> Nucleotide sequence encoding apolipoprotein
C-III(APOC3)

<400> 25
tgctcagttc atccctagag gcagctgctc caggaacaga ggtgcc atg cag ccc 55

```
                                                         Met Gln Pro
                                                          1

cgg gta ctc ctt gtt gtt gcc ctc ctg gcg ctc ctg gcc tct gcc cga      103
Arg Val Leu Leu Val Val Ala Leu Leu Ala Leu Leu Ala Ser Ala Arg
     5                  10                  15

gct tca gag gcc gag gat gcc tcc ctt ctc agc ttc atg cag ggt tac      151
Ala Ser Glu Ala Glu Asp Ala Ser Leu Leu Ser Phe Met Gln Gly Tyr
 20                  25                  30                  35

atg aag cac gcc acc aag acc gcc aag gat gca ctg agc agc gtg cag      199
Met Lys His Ala Thr Lys Thr Ala Lys Asp Ala Leu Ser Ser Val Gln
                 40                  45                  50

gag tcc cag gtg gcc cag cag gcc agg ggc tgg gtg acc gat ggc ttc      247
Glu Ser Gln Val Ala Gln Gln Ala Arg Gly Trp Val Thr Asp Gly Phe
             55                  60                  65

agt tcc ctg aaa gac tac tgg agc acc gtt aag gac aag ttc tct gag      295
Ser Ser Leu Lys Asp Tyr Trp Ser Thr Val Lys Asp Lys Phe Ser Glu
         70                  75                  80

ttc tgg gat ttg gac cct gag gtc aga cca act tca gcc gtg gct gcc      343
Phe Trp Asp Leu Asp Pro Glu Val Arg Pro Thr Ser Ala Val Ala Ala
     85                  90                  95

tga gacctcaata ccccaagtcc acctgcctat ccatcctgcg agctccttgg           396
 *


gtcctgcaat ctccagggct gcccctgtag gttgcttaaa agggacagta ttctcagtgc    456
tctcctaccc cacctcatgc ctggcccccc tccaggcatg ctggcctccc aataaagctg    516
gacaagaagc tgctatg                                                   533

<210> 26
<211> 99
<212> PRT
<213> Homo sapien

<400> 26
Met Gln Pro Arg Val Leu Leu Val Val Ala Leu Leu Ala Leu Leu Ala
 1               5                  10                  15
Ser Ala Arg Ala Ser Glu Ala Glu Asp Ala Ser Leu Leu Ser Phe Met
            20                  25                  30
Gln Gly Tyr Met Lys His Ala Thr Lys Thr Ala Lys Asp Ala Leu Ser
        35                  40                  45
Ser Val Gln Glu Ser Gln Val Ala Gln Gln Ala Arg Gly Trp Val Thr
    50                  55                  60
Asp Gly Phe Ser Ser Leu Lys Asp Tyr Trp Ser Thr Val Lys Asp Lys
65                  70                  75                  80
Phe Ser Glu Phe Trp Asp Leu Asp Pro Glu Val Arg Pro Thr Ser Ala
                85                  90                  95
Val Ala Ala


<210> 27
<211> 8925
```

```
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (5020)...(6162)
<223> Nucleotide encoding ATP-binding cassette (ABC1)

<223> n= a or g or c or t

<400> 27
ctcagtgtca gctgctgctg gaagtggcct ggcctctatt tatcttcctg atcctgatct      60
ctgttcggct gagctaccca ccctatgaac aacatgaatg ccattttcca aataaagcca     120
tgccctctgc aggaacactt ccttgggttc aggggattat ctgtaatgcc aacaacccct     180
gtttccgtta cccgactcct ggggaggctc ccggagttgt tggaaacttt aacaaatcca     240
ttgtggctcg cctgttctca gatgctcgga ggcttctttt atacagccag aaagacacca     300
gcatgaagga catgcgcaaa gttctgagaa cattacagca gatcaagaaa tccagctcaa     360
acttgaagct tcaagatttc ctggtggaca atgaaacctt ctctgggttc ctgtatcaca     420
acctctctct cccaaagtct actgtggaca agatgctgag ggctgatgtc attctccaca     480
aggtattttt gcaaggctac cagttacatt tgacaagtct gtgcaatgga tcaaaatcag     540
aagagatgat tcaacttggt gaccaagaag tttctgagct ttgtggccta ccaagggaga     600
aactggctgc agcagagcga gtacttcgtt ccaacatgga catcctgaag ccaatcctga     660
gaacactaaa ctctacatct cccttcccga gcaaggagct ggctgaagcc acaaaaacat     720
tgctgcatag tcttgggact ctggcccagg agctgttcag catgagaagc tggagtgaca     780
tgcgacagga ggtgatgttt ctgaccaatg tgaacagctc cagctcctcc acccaaatct     840
accaggctgt gtctcgtatt gtctgcgggc atcccgaggg aggggggctg aagatcaagt     900
ctctcaactg gtatgaggac aacaactaca aagccctctt tggaggcaat ggcactgagg     960
aagatgctga aaccttctat gacaactcta caactcctta ctgcaatgat ttgatgaaga    1020
atttggagtc tagtcctctt tcccgcatta tctggaaagc tctgaagccg ctgctcgttg    1080
ggaagatcct gtatacacct gacactccag ccacaaggca ggtcatggct gaggtgaaca    1140
agaccttcca ggaactggct gtgttccatg atctggaagg catgtgggag gaactcagcc    1200
ccaagatctg gaccttcatg gagaacagcc aagaaatgga ccttgtccgg atgctgttgg    1260
acagcaggga caatgaccac ttttgggaac agcagttgga tggcttagat tggacagccc    1320
aagacatcgt ggcgtttttg gccaagcacc cagaggatgt ccagtccagt aatggttctg    1380
tgtacacctg gagagaagct ttcaacgaga ctaaccaggc aatccggacc atatctcgct    1440
tcatggagtg tgtcaacctg aacaagctag aacccatagc aacagaagtc tggctcatca    1500
acaagtccat ggagctgctg gatgagagga agttctgggc tggtattgtg ttcactggaa    1560
ttactccagg cagcattgag ctgccccatc atgtcaagta caagatccga atggacattg    1620
acaatgtgga gaggacaaat aaaatcaagg atgggtactg ggaccctggt cctcgagctg    1680
accccttga ggacatgcgg tacgtctggg ggggcttcgc ctacttgcag gatgtggtgg    1740
agcaggcaat catcagggtg ctgacgggca ccgagaagaa aactggtgtc tatatgcaac    1800
agatgcccta tccctgttac gttgatgaca tctttctgcg ggtgatgagc cggtcaatgc    1860
ccctcttcat gacgctggcc tggatttact cagtggctgt gatcatcaag ggcatcgtgt    1920
atgagaagga ggcacggctg aaagagacca tgcggatcat gggcctggac aacagcatcc    1980
tctggtttag ctggttcatt agtagcctca ttcctcttct tgtgagcgct ggcctgctag    2040
tggtcatcct gaagttagga aacctgctgc cctacagtga tcccagcgtg gtgtttgtct    2100
tcctgtccgt gtttgctgtg gtgacaatcc tgcagtgctt cctgattagc acactcttct    2160
ccagagccaa cctggcagca gcctgtgggg gcatcatcta cttcacgctg tacctgccct    2220
acgtcctgtg tgtggcatgg caggactacg tgggcttcac actcaagatc ttcgctagcc    2280
tgctgtctcc tgtggctttt gggtttggct gtgagtactt tgcccttttt gaggagcagg    2340
gcattggagt gcagtgggac aacctgtttg agagtcctgt ggaggaagat ggcttcaatc    2400
tcaccacttc ggtctccatg atgctgtttg acaccttcct ctatggggtg atgacctggt    2460
acattgaggc tgtctttcca ggccagtacg gaattcccag gccctggtat tttccttgca    2520
ccaagtccta ctggtttggc gaggaaagtg atgagaagag ccaccctggt tccaaccaga    2580
agagaatatc agaaatctgc atggaggagg aacccaccca cttgaagctg ggcgtgtcca    2640
ttcagaacct ggtaaaagtc taccgagatg ggatgaaggt ggctgtcgat ggcctggcac    2700
tgaatttta tgagggccag atcacctcct tcctgggcca caatggagcg gggaagacga    2760
ccaccatgtc aatcctgacc gggttgttcc ccccgacctc gggcaccgcc tacatcctgg    2820
```

```
gaaaagacat tcgctctgag atgagcacca tccggcagaa cctgggggtc tgtccccagc   2880
ataacgtgct gtttgacatg ctgactgtcg aagaacacat ctggttctat gcccgcttga   2940
aagggctctc tgagaagcac gtgaaggcgg agatggagca gatggccctg gatgttggtt   3000
tgccatcaag caagctgaaa agcaaaacaa gccagctgtc aggtggaatg cagagaaagc   3060
tatctgtggc cttggccttt gtcgggggat ctaaggttgt cattctggat gaacccacag   3120
ctggtgtgga cccttactcc cgcaggggaa tatgggagct gctgctgaaa taccgacaag   3180
gccgcaccat tattctctct acacaccaca tggatgaagc ggacgtcctg gggggacagga  3240
ttgccatcat ctcccatggg aagctgtgct gtgtgggctc ctccctgttt ctgaagaacc   3300
agctgggaac aggctactac ctgaccttgg tcaagaaaga tgtggaatcc tccctcagtt   3360
cctgcagaaa cagtagtagc actgtgtcat acctgaaaaa ggaggacagt gtttctcaga   3420
gcagttctga tgctggcctg ggcagcgacc atgagagtga cacgctgacc atcgatgtct   3480
ctgctatctc caacctcatc aggaagcatg tgtctgaagc ccggctggtg gaagacatag   3540
ggcatgagct gacctatgtg ctgccatatg aagctgctaa ggagggagcc tttgtggaac   3600
tctttcatga gattgatgac cggctctcag acctgggcat ttctagttat ggcatctcag   3660
agacgaccct ggaagaaata ttcctcaagg tggccgaaga gagtggggtg gatgctgaga   3720
cctcagatgg taccttgcca gcaagacgaa acaggcgggc cttcggggac aagcagagct   3780
gtcttcgccc gttcactgaa gatgatgctg ctgatccaaa tgattctgac atagacccag   3840
aatccagaga gacagacttg ctcagtggga tggatggcaa agggtcctac caggtgaaag   3900
gctggaaact tacacagcaa cagtttgtgg ccctttttgtg gaagagactg ctaattgcca  3960
gacggagtcg gaaaggattt tttgctcaga ttgtcttgcc agctgtgttt gtctgcattg   4020
cccttgtgtt cagcctgatc gtgccacccct ttggcaagta ccccagcctg gaacttcagc  4080
cctggatgta caacgaacag tacacatttg tcagcaatga tgctcctgag gacacgggaa   4140
ccctggaact cttaaacgcc ctcaccaaag accctggctt cgggacccgc tgtatggaag   4200
gaaacccaat cccagacacg ccctgccagg caggggagga agagtggacc actgccccag   4260
ttccccagac catcatggac ctcttccaga atgggaactg gacaatgcag aacccttcac   4320
ctgcatgcca gtgtagcagc gacaaaatca agaagatgct gcctgtgtgt cccccagggg   4380
caggggggct gcctcctcca caaagaaaac aaaacactgc agatatcctt caggacctga   4440
caggaagaaa catttcggat tatctggtga agacgtatgt gcagatcata gccaaaagct   4500
taaagaacaa gatctgggtg aatgagttta ggtatggcgg cttttccctg ggtgtcagta   4560
atactcaagc acttcctccg agtcaagaag ttaatgatgc catcaaacaa atgaagaaac   4620
acctaaagct ggccaaggac agttctgcag atcgatttct caacagcttg ggaagattta   4680
tgacaggact ggacaccaaa aataatgtca aggtgtggtt caataacaag ggctggcatg   4740
caatcagctc tttcctgaat gtcatcaaca atgccattct ccgggccaac ctgcaaaagg   4800
gagagaaccc tagccattat ggaattactg ctttcaatca tcccctgaat ctcaccaagc   4860
agcagctctc agaggtggct cggatgacca catcagtgga tgtccttgtg tccatctgtg   4920
tcatctttgc aatgtccttc gtcccagcca gctttgtcgt attcctgatc caggagcggg   4980
tcagcaaagc aaaacacctg cagttcatca gtggagtga agc ctg tca tct act      5034
                                           Ser Leu Ser Ser Thr
                                           1               5

ggc tct cta att ttg tct ggg ata tgt gca att aag ttg ttt cca ann      5082
Gly Ser Leu Ile Leu Ser Gly Ile Cys Ala Ile Lys Leu Phe Pro Xaa
                10                  15                  20

nnn nnn nnn nnn nnn nnn nnn nnn nnn nnn nnn nnn nnn nnn nnn nnn      5130
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
            25                  30                  35

nnn nnn nnn nnn nnn nnn nnn nnn nnn nnn nnn nnn nnn nnn nnn nnn      5178
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
        40                  45                  50

nta atc ttt cct ttt cag tgc ttt ggg ctc ctg gga gtt aat ggg gct      5226
Xaa Ile Phe Pro Phe Gln Cys Phe Gly Leu Leu Gly Val Asn Gly Ala
    55                  60                  65

gga aaa tca tca act ttc aag atg tta aca gga gat acc act gtt acc      5274
Gly Lys Ser Ser Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr
```

```
 70                  75                  80                  85

aga gga gat gct ttc ctt aac att tgc agt atc tta tca aac atc cat      5322
Arg Gly Asp Ala Phe Leu Asn Ile Cys Ser Ile Leu Ser Asn Ile His
                 90                  95                 100

gaa gta cat cag aac atg ggc tac tgc cct cag ttt gat gcc atc aca      5370
Glu Val His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Thr
            105                 110                 115

gag ctg ttg act ggg aga gaa cac gtg gag ttc ttt gcc ctt ttg aga      5418
Glu Leu Leu Thr Gly Arg Glu His Val Glu Phe Phe Ala Leu Leu Arg
        120                 125                 130

gga gtc cca gag aaa gaa gtt ggc aag gtt ggt gag tgg gcg att cgg      5466
Gly Val Pro Glu Lys Glu Val Gly Lys Val Gly Glu Trp Ala Ile Arg
    135                 140                 145

aaa ctg ggc ctc gtg aag tat gga gaa aaa tat gct ggt aac tat agt      5514
Lys Leu Gly Leu Val Lys Tyr Gly Glu Lys Tyr Ala Gly Asn Tyr Ser
150                 155                 160                 165

gga ggc aac aaa cgc aag ctc tct aca gcc atg gct ttg atc ggc ggg      5562
Gly Gly Asn Lys Arg Lys Leu Ser Thr Ala Met Ala Leu Ile Gly Gly
                170                 175                 180

cct cct gtg gtg ttt ctg gat gaa ccc acc aca ggc atg gat ccc aaa      5610
Pro Pro Val Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Lys
            185                 190                 195

gcc cgg cgg ttc ttg tgg aat tgt gcc cta agt gtt gtc aag gag ggg      5658
Ala Arg Arg Phe Leu Trp Asn Cys Ala Leu Ser Val Val Lys Glu Gly
        200                 205                 210

aga tca gta gtg ctt aca tct cat agt atg gaa gaa tgt gaa gct ctt      5706
Arg Ser Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu
    215                 220                 225

tgc act agg atg gca atc atg gtc aat gga agg ttc agg tgc ctt ggc      5754
Cys Thr Arg Met Ala Ile Met Val Asn Gly Arg Phe Arg Cys Leu Gly
230                 235                 240                 245

agt gtc cag cat cta aaa aat agg ttt gga gat ggt tat aca ata gtt      5802
Ser Val Gln His Leu Lys Asn Arg Phe Gly Asp Gly Tyr Thr Ile Val
                250                 255                 260

gta cga ata gca ggg tcc aac ccg gac ctg aag cct gtc cag gat ttc      5850
Val Arg Ile Ala Gly Ser Asn Pro Asp Leu Lys Pro Val Gln Asp Phe
            265                 270                 275

ttt gga ctt gca ttt cct gga agt gtt cta aaa gag aaa cac cgg aac      5898
Phe Gly Leu Ala Phe Pro Gly Ser Val Leu Lys Glu Lys His Arg Asn
        280                 285                 290

atg cta caa tac cag ctt cca tct tca tta tct tct ctg gcc agg ata      5946
Met Leu Gln Tyr Gln Leu Pro Ser Ser Leu Ser Ser Leu Ala Arg Ile
    295                 300                 305
```

```
ttc agc atc ctc tcc cag agc aaa aag cga ctc cac ata gaa gac tac      5994
Phe Ser Ile Leu Ser Gln Ser Lys Lys Arg Leu His Ile Glu Asp Tyr
310                 315                 320                 325

tct gtt tct cag aca aca ctt gac caa gta ttt gtg aac ttt gcc aag      6042
Ser Val Ser Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys
                330                 335                 340

gac caa agt gat gat gac cac tta aaa gac ctc tca tta cac aaa aac      6090
Asp Gln Ser Asp Asp Asp His Leu Lys Asp Leu Ser Leu His Lys Asn
            345                 350                 355

cag aca gta gtg gac gtt gca gtt ctc aca tct ttt cta cag gat gag      6138
Gln Thr Val Val Asp Val Ala Val Leu Thr Ser Phe Leu Gln Asp Glu
        360                 365                 370

aaa gtg aaa gaa agc tat gta tga agaatcctgt tcatacgggg tggctgaaag     6192
Lys Val Lys Glu Ser Tyr Val  *
    375                 380

taaagaggaa ctagactttc ctttgcacca tgtgaagtgt tgtggagaaa agagccagaa    6252
gttgatgtgg gaagaagtaa actggatact gtactgatac tattcaatgc aatgcaattc    6312
aatgcaatga aaacaaaatt ccattacagg ggcagtgcct ttgtagccta tgtcttgtat    6372
ggctctcaag tgaaagactt gaatttagtt ttttacctat acctatgtga aactctatta    6432
tggaacccaa tggacatatg ggtttgaact cacacttttt tttttttttt tgttcctgtg    6492
tattctcatt ggggttgcaa caataattca tcaagtaatc atggccagcg attattgatc    6552
aaaatcaaaa ggtaatgcac atcctcattc actaagccat gccatgccca ggagactggt    6612
ttcccggtga cacatccatt gctggcaatg agtgtgccag agttattagt gccaagtttt    6672
tcagaaagtt tgaagcacca tggtgtgtca tgctcacttt tgtgaaagct gctctgctca    6732
gagtctatca acattgaata tcagttgaca gaatggtgcc atgcgtggct aacatcctgc    6792
tttgattccc tctgataagc tgttctggtg gcagtaacat gcaacaaaaa tgtgggtgtc    6852
tccaggcacg ggaaacttgg ttccattgtt atattgtcct atgcttcgag ccatgggtct    6912
acagggtcat ccttatgaga ctcttaaata tacttagatc ctggtaagag gcaaagaatc    6972
aacagccaaa ctgctggggc tgcaagctgc tgaagccagg gcatgggatt aaagagattg    7032
tgcgttcaaa cctagggaag cctgtgccca tttgtcctga ctgtctgcta acatggtaca    7092
ctgcatctca agatgtttat ctgacacaag tgtattattt ctggcttttt gaattaatct    7152
agaaaatgaa aagatggagt tgtattttga caaaaatgtt tgtacttttt aatgttattt    7212
ggaattttaa gttctatcag tgacttctga atccttagaa tggcctcttt gtagaaccct    7272
gtggtataga ggagtatggc cactgcccca ctatttttat tttcttatgt aagtttgcat    7332
atcagtcatg actagtgcct agaaagcaat gtgatggtca ggatctcatg acattatatt    7392
tgagtttctt tcagatcatt taggatactc ttaatctcac ttcatcaatc aaatattttt    7452
tgagtgtatg ctgtagctga aagagtatgt acgtacgtat aagactagag agatattaag    7512
tctcagtaca cttcctgtgc catgttattc agctcactgg tttacaaata taggttgtct    7572
tgtggttgta ggagcccact gtaacaatac tgggcagcct tttttttttt tttttttaatt    7632
gcaacaatgc aaaagccaag aaagtataag ggtcacaagt ctaaacaatg aattcttcaa    7692
cagggaaaac agctagcttg aaaacttgct gaaaaacaca acttgtgttt atggcattta    7752
gtaccttcaa ataattggct ttgcagatat tggatacccc attaaatctg acagtctcaa    7812
atttttcatc tcttcaatca ctagtcaaga aaaatataaa aacaacaaat acttccatat    7872
ggagcatttt tcagagtttt ctaacccagt cttatttttc tagtcagtaa acatttgtaa    7932
aaatactgtt tcactaatac ttactgttaa ctgtcttgag agaaaagaaa aatatgagag    7992
aactattgtt tggggaagtt caagtgatct ttcaatatca ttactaactt cttccacttt    8052
ttccagaatt tgaatattaa cgctaaaggt gtaagacttc agatttcaaa ttaatctttc    8112
tatatttttt aaatttacag aatattatat aacccactgc tgaaaaagaa aaaaatgatt    8172
gttttagaag ttaaagtcaa tattgatttt aaatataagt aatgaaggca tatttccaat    8232
aactagtgat atggcatcgt tgcattttac agtatcttca aaaatacaga atttatagaa    8292
taatttctcc tcatttaata tttttcaaaa tcaaagttat ggtttcctca ttttactaaa    8352
atcgtattct aattcttcat tatagtaaat ctatgagcaa ctccttactt cggttcctct    8412
gatttcaagg ccatatttta aaaaatcaaa aggcactgtg aactattttg aagaaaacac    8472
```

```
aacattttaa tacagattga aaggacctct tctgaagcta gaaacaatct atagttatac   8532
atcttcatta atactgtgtt accttttaaa atagtaattt tttacatttt cctgtgtaaa   8592
cctaattgtg gtagaaattt ttaccaactc tatactcaat caagcaaaat ttctgtatat   8652
tccctgtgga atgtacctat gtgagtttca gaaattctca aaatacgtgt tcaaaaattt   8712
ctgcttttgc atctttggga cacctcagaa aacttattaa caactgtgaa tatgagaaat   8772
acagaagaaa ataataagcc ctctatacat aaatgcccag cacaattcat tgttaaaaaa   8832
caaccaaacc tcacactact gtatttcatt atctgtactg aaagcaaatg ctttgtgact   8892
attaaatgtt gcacatcatt cattcactgt ata                                 8925
```

<210> 28
<211> 380
<212> PRT
<213> Homo sapien

<220>
<221> UNSURE
<222> (21)...(54)
<223> Xaa = unknown

<400> 28

```
Ser Leu Ser Ser Thr Gly Ser Leu Ile Leu Ser Gly Ile Cys Ala Ile
 1               5                  10                  15
Lys Leu Phe Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
            20                  25                  30
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
        35                  40                  45
Xaa Xaa Xaa Xaa Xaa Xaa Ile Phe Pro Phe Gln Cys Phe Gly Leu Leu
    50                  55                  60
Gly Val Asn Gly Ala Gly Lys Ser Ser Thr Phe Lys Met Leu Thr Gly
65                  70                  75                  80
Asp Thr Thr Val Thr Arg Gly Asp Ala Phe Leu Asn Ile Cys Ser Ile
                85                  90                  95
Leu Ser Asn Ile His Glu Val His Gln Asn Met Gly Tyr Cys Pro Gln
            100                 105                 110
Phe Asp Ala Ile Thr Glu Leu Leu Thr Gly Arg Glu His Val Glu Phe
        115                 120                 125
Phe Ala Leu Leu Arg Gly Val Pro Glu Lys Glu Val Gly Lys Val Gly
    130                 135                 140
Glu Trp Ala Ile Arg Lys Leu Gly Leu Val Lys Tyr Gly Glu Lys Tyr
145                 150                 155                 160
Ala Gly Asn Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ser Thr Ala Met
                165                 170                 175
Ala Leu Ile Gly Gly Pro Pro Val Val Phe Leu Asp Glu Pro Thr Thr
            180                 185                 190
Gly Met Asp Pro Lys Ala Arg Arg Phe Leu Trp Asn Cys Ala Leu Ser
        195                 200                 205
Val Val Lys Glu Gly Arg Ser Val Val Leu Thr Ser His Ser Met Glu
    210                 215                 220
Glu Cys Glu Ala Leu Cys Thr Arg Met Ala Ile Met Val Asn Gly Arg
225                 230                 235                 240
Phe Arg Cys Leu Gly Ser Val Gln His Leu Lys Asn Arg Phe Gly Asp
                245                 250                 255
Gly Tyr Thr Ile Val Val Arg Ile Ala Gly Ser Asn Pro Asp Leu Lys
            260                 265                 270
Pro Val Gln Asp Phe Phe Gly Leu Ala Phe Pro Gly Ser Val Leu Lys
        275                 280                 285
Glu Lys His Arg Asn Met Leu Gln Tyr Gln Leu Pro Ser Ser Leu Ser
    290                 295                 300
```

Ser Leu Ala Arg Ile Phe Ser Ile Leu Ser Gln Ser Lys Lys Arg Leu
305 310 315 320
His Ile Glu Asp Tyr Ser Val Ser Gln Thr Thr Leu Asp Gln Val Phe
325 330 335
Val Asn Phe Ala Lys Asp Gln Ser Asp Asp Asp His Leu Lys Asp Leu
340 345 350
Ser Leu His Lys Asn Gln Thr Val Val Asp Val Ala Val Leu Thr Ser
355 360 365
Phe Leu Gln Asp Glu Lys Val Lys Glu Ser Tyr Val
370 375 380

<210> 29
<211> 897
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (39)...(842)
<223> Nucleotide sequence encoding apolipoprotein A-1
(APOA1)

<400> 29
agagactgcg agaaggaggt cccccacggc ccttcagg atg aaa gct gcg gtg ctg 56
Met Lys Ala Ala Val Leu
1 5

acc ttg gcc gtg ctc ttc ctg acg ggg agc cag gct cgg cat ttc tgg 104
Thr Leu Ala Val Leu Phe Leu Thr Gly Ser Gln Ala Arg His Phe Trp
10 15 20

cag caa gat gaa ccc ccc cag agc ccc tgg gat cga gtg aag gac ctg 152
Gln Gln Asp Glu Pro Pro Gln Ser Pro Trp Asp Arg Val Lys Asp Leu
25 30 35

gcc act gtg tac gtg gat gtg ctc aaa gac agc ggc aga gac tat gtg 200
Ala Thr Val Tyr Val Asp Val Leu Lys Asp Ser Gly Arg Asp Tyr Val
40 45 50

tcc cag ttt gaa ggc tcc gcc ttg gga aaa cag cta aac cta aag ctc 248
Ser Gln Phe Glu Gly Ser Ala Leu Gly Lys Gln Leu Asn Leu Lys Leu
55 60 65 70

ctt gac aac tgg gac agc gtg acc tcc acc ttc agc aag ctg cgc gaa 296
Leu Asp Asn Trp Asp Ser Val Thr Ser Thr Phe Ser Lys Leu Arg Glu
75 80 85

cag ctc ggc cct gtg acc cag gag ttc tgg gat aac ctg gaa aag gag 344
Gln Leu Gly Pro Val Thr Gln Glu Phe Trp Asp Asn Leu Glu Lys Glu
90 95 100

aca gag ggc ctg agg cag gag atg agc aag gat ctg gag gag gtg aag 392
Thr Glu Gly Leu Arg Gln Glu Met Ser Lys Asp Leu Glu Glu Val Lys
105 110 115

gcc aag gtg cag ccc tac ctg gac gac ttc cag aag aag tgg cag gag 440
Ala Lys Val Gln Pro Tyr Leu Asp Asp Phe Gln Lys Lys Trp Gln Glu
120 125 130

```
gag atg gag ctc tac cgc cag aag gtg gag ccg ctg cgc gca gag ctc        488
Glu Met Glu Leu Tyr Arg Gln Lys Val Glu Pro Leu Arg Ala Glu Leu
135                 140                 145                 150

caa gag ggc gcg cgc cag aag ctg cac gag ctg caa gag aag ctg agc        536
Gln Glu Gly Ala Arg Gln Lys Leu His Glu Leu Gln Glu Lys Leu Ser
                155                 160                 165

cca ctg ggc gag gag atg cgc gac cgc gcg cgc gcc cat gtg gac gcg        584
Pro Leu Gly Glu Glu Met Arg Asp Arg Ala Arg Ala His Val Asp Ala
            170                 175                 180

ctg cgc acg cat ctg gcc ccc tac agc gac gag ctg cgc cag cgc ttg        632
Leu Arg Thr His Leu Ala Pro Tyr Ser Asp Glu Leu Arg Gln Arg Leu
        185                 190                 195

gcc gcg cgc ctt gag gct ctc aag gag aac ggc ggc gcc aga ctg gcc        680
Ala Ala Arg Leu Glu Ala Leu Lys Glu Asn Gly Gly Ala Arg Leu Ala
    200                 205                 210

gag tac cac gcc aag gcc acc gag cat ctg agc acg ctc agc gag aag        728
Glu Tyr His Ala Lys Ala Thr Glu His Leu Ser Thr Leu Ser Glu Lys
215                 220                 225                 230

gcc aag ccc gcg ctc gag gac ctc cgc caa ggc ctg ctg ccc gtg ctg        776
Ala Lys Pro Ala Leu Glu Asp Leu Arg Gln Gly Leu Leu Pro Val Leu
                235                 240                 245

gag agc ttc aag gtc agc ttc ctg agc gct ctc gag gag tac act aag        824
Glu Ser Phe Lys Val Ser Phe Leu Ser Ala Leu Glu Glu Tyr Thr Lys
            250                 255                 260

aag ctc aac acc cag tga ggcgcccgcc gccgccccc ttcccggtgc                872
Lys Leu Asn Thr Gln  *
        265

tcagaataaa cgtttccaaa gtggg                                            897

<210> 30
<211> 267
<212> PRT
<213> Homo sapien

<400> 30
Met Lys Ala Ala Val Leu Thr Leu Ala Val Leu Phe Leu Thr Gly Ser
 1               5                  10                  15
Gln Ala Arg His Phe Trp Gln Gln Asp Glu Pro Pro Gln Ser Pro Trp
            20                  25                  30
Asp Arg Val Lys Asp Leu Ala Thr Val Tyr Val Asp Val Leu Lys Asp
        35                  40                  45
Ser Gly Arg Asp Tyr Val Ser Gln Phe Glu Gly Ser Ala Leu Gly Lys
    50                  55                  60
Gln Leu Asn Leu Lys Leu Leu Asp Asn Trp Asp Ser Val Thr Ser Thr
65                  70                  75                  80
Phe Ser Lys Leu Arg Glu Gln Leu Gly Pro Val Thr Gln Glu Phe Trp
                85                  90                  95
Asp Asn Leu Glu Lys Glu Thr Glu Gly Leu Arg Gln Glu Met Ser Lys
```

```
            100                     105                     110
Asp Leu Glu Glu Val Lys Ala Lys Val Gln Pro Tyr Leu Asp Asp Phe
        115                     120                     125
Gln Lys Lys Trp Gln Glu Glu Met Glu Leu Tyr Arg Gln Lys Val Glu
    130                     135                     140
Pro Leu Arg Ala Glu Leu Gln Glu Gly Ala Arg Gln Lys Leu His Glu
145                     150                     155                 160
Leu Gln Glu Lys Leu Ser Pro Leu Gly Glu Glu Met Arg Asp Arg Ala
                165                     170                     175
Arg Ala His Val Asp Ala Leu Arg Thr His Leu Ala Pro Tyr Ser Asp
            180                     185                     190
Glu Leu Arg Gln Arg Leu Ala Ala Arg Leu Glu Ala Leu Lys Glu Asn
        195                     200                     205
Gly Gly Ala Arg Leu Ala Glu Tyr His Ala Lys Ala Thr Glu His Leu
    210                     215                     220
Ser Thr Leu Ser Glu Lys Ala Lys Pro Ala Leu Glu Asp Leu Arg Gln
225                     230                     235                 240
Gly Leu Leu Pro Val Leu Glu Ser Phe Lys Val Ser Phe Leu Ser Ala
                245                     250                     255
Leu Glu Glu Tyr Thr Lys Lys Leu Asn Thr Gln
            260                     265

<210> 31
<211> 14121
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (129)...(13820)
<223> Nucleotide sequence encoding apolipoprotein B
      (APOB)

<400> 31
attcccaccg ggacctgcgg ggctgagtgc ccttctcggt tgctgccgct gaggagcccg      60
cccagccagc cagggccgcg aggccgaggc caggccgcag cccaggagcc gccccaccgc     120
agctggcg atg gac ccg ccg agg ccc gcg ctg ctg gcg ctg ctg gcg ctg      170
         Met Asp Pro Pro Arg Pro Ala Leu Leu Ala Leu Leu Ala Leu
          1               5                  10

cct gcg ctg ctg ctg ctg ctg ctg gcg ggc gcc agg gcc gaa gag gaa      218
Pro Ala Leu Leu Leu Leu Leu Leu Ala Gly Ala Arg Ala Glu Glu Glu
 15                  20                  25                  30

atg ctg gaa aat gtc agc ctg gtc tgt cca aaa gat gcg acc cga ttc      266
Met Leu Glu Asn Val Ser Leu Val Cys Pro Lys Asp Ala Thr Arg Phe
                 35                  40                  45

aag cac ctc cgg aag tac aca tac aac tat gag gct gag agt tcc agt      314
Lys His Leu Arg Lys Tyr Thr Tyr Asn Tyr Glu Ala Glu Ser Ser Ser
             50                  55                  60

gga gtc cct ggg act gct gat tca aga agt gcc acc agg atc aac tgc      362
Gly Val Pro Gly Thr Ala Asp Ser Arg Ser Ala Thr Arg Ile Asn Cys
         65                  70                  75

aag gtt gag ctg gag gtt ccc cag ctc tgc agc ttc atc ctg aag acc      410
Lys Val Glu Leu Glu Val Pro Gln Leu Cys Ser Phe Ile Leu Lys Thr
```

```
     80                  85                  90

agc cag tgc acc ctg aaa gag gtg tat ggc ttc aac cct gag ggc aaa       458
Ser Gln Cys Thr Leu Lys Glu Val Tyr Gly Phe Asn Pro Glu Gly Lys
 95                 100                 105                 110

gcc ttg ctg aag aaa acc aag aac tct gag gag ttt gct gca gcc atg       506
Ala Leu Leu Lys Lys Thr Lys Asn Ser Glu Glu Phe Ala Ala Ala Met
                115                 120                 125

tcc agg tat gag ctc aag ctg gcc att cca gaa ggg aag cag gtt ttc       554
Ser Arg Tyr Glu Leu Lys Leu Ala Ile Pro Glu Gly Lys Gln Val Phe
            130                 135                 140

ctt tac ccg gag aaa gat gaa cct act tac atc ctg aac atc aag agg       602
Leu Tyr Pro Glu Lys Asp Glu Pro Thr Tyr Ile Leu Asn Ile Lys Arg
        145                 150                 155

ggc atc att tct gcc ctc ctg gtt ccc cca gag aca gaa gaa gcc aag       650
Gly Ile Ile Ser Ala Leu Leu Val Pro Pro Glu Thr Glu Glu Ala Lys
    160                 165                 170

caa gtg ttg ttt ctg gat acc gtg tat gga aac tgc tcc act cac ttt       698
Gln Val Leu Phe Leu Asp Thr Val Tyr Gly Asn Cys Ser Thr His Phe
175                 180                 185                 190

acc gtc aag acg agg aag ggc aat gtg gca aca gaa ata tcc act gaa       746
Thr Val Lys Thr Arg Lys Gly Asn Val Ala Thr Glu Ile Ser Thr Glu
                195                 200                 205

aga gac ctg ggg cag tgt gat cgc ttc aag ccc atc cgc aca ggc atc       794
Arg Asp Leu Gly Gln Cys Asp Arg Phe Lys Pro Ile Arg Thr Gly Ile
            210                 215                 220

agc cca ctt gct ctc atc aaa ggc atg acc cgc ccc ttg tca act ctg       842
Ser Pro Leu Ala Leu Ile Lys Gly Met Thr Arg Pro Leu Ser Thr Leu
        225                 230                 235

atc agc agc agc cag tcc tgt cag tac aca ctg gac gct aag agg aag       890
Ile Ser Ser Ser Gln Ser Cys Gln Tyr Thr Leu Asp Ala Lys Arg Lys
    240                 245                 250

cat gtg gca gaa gcc atc tgc aag gag caa cac ctc ttc ctg cct ttc       938
His Val Ala Glu Ala Ile Cys Lys Glu Gln His Leu Phe Leu Pro Phe
255                 260                 265                 270

tcc tac aac aat aag tat ggg atg gta gca caa gtg aca cag act ttg       986
Ser Tyr Asn Asn Lys Tyr Gly Met Val Ala Gln Val Thr Gln Thr Leu
                275                 280                 285

aaa ctt gaa gac aca cca aag atc aac agc cgc ttc ttt ggt gaa ggt      1034
Lys Leu Glu Asp Thr Pro Lys Ile Asn Ser Arg Phe Phe Gly Glu Gly
            290                 295                 300

act aag aag atg ggc ctc gca ttt gag agc acc aaa tcc aca tca cct      1082
Thr Lys Lys Met Gly Leu Ala Phe Glu Ser Thr Lys Ser Thr Ser Pro
        305                 310                 315
```

```
cca aag cag gcc gaa gct gtt ttg aag act ctc cag gaa ctg aaa aaa      1130
Pro Lys Gln Ala Glu Ala Val Leu Lys Thr Leu Gln Glu Leu Lys Lys
    320                 325                 330

cta acc atc tct gag caa aat atc cag aga gct aat ctc ttc aat aag      1178
Leu Thr Ile Ser Glu Gln Asn Ile Gln Arg Ala Asn Leu Phe Asn Lys
335                 340                 345                 350

ctg gtt act gag ctg aga ggc ctc agt gat gaa gca gtc aca tct ctc      1226
Leu Val Thr Glu Leu Arg Gly Leu Ser Asp Glu Ala Val Thr Ser Leu
                355                 360                 365

ttg cca cag ctg att gag gtg tcc agc ccc atc act tta caa gcc ttg      1274
Leu Pro Gln Leu Ile Glu Val Ser Ser Pro Ile Thr Leu Gln Ala Leu
            370                 375                 380

gtt cag tgt gga cag cct cag tgc tcc act cac atc ctc cag tgg ctg      1322
Val Gln Cys Gly Gln Pro Gln Cys Ser Thr His Ile Leu Gln Trp Leu
        385                 390                 395

aaa cgt gtg cat gcc aac ccc ctt ctg ata gat gtg gtc acc tac ctg      1370
Lys Arg Val His Ala Asn Pro Leu Leu Ile Asp Val Val Thr Tyr Leu
    400                 405                 410

gtg gcc ctg atc ccc gag ccc tca gca cag cag ctg cga gag atc ttc      1418
Val Ala Leu Ile Pro Glu Pro Ser Ala Gln Gln Leu Arg Glu Ile Phe
415                 420                 425                 430

aac atg gcg agg gat cag cgc agc cga gcc acc ttg tat gcg ctg agc      1466
Asn Met Ala Arg Asp Gln Arg Ser Arg Ala Thr Leu Tyr Ala Leu Ser
                435                 440                 445

cac gcg gtc aac aac tat cat aag aca aac cct aca ggg acc cag gag      1514
His Ala Val Asn Asn Tyr His Lys Thr Asn Pro Thr Gly Thr Gln Glu
            450                 455                 460

ctg ctg gac att gct aat tac ctg atg gaa cag att caa gat gac tgc      1562
Leu Leu Asp Ile Ala Asn Tyr Leu Met Glu Gln Ile Gln Asp Asp Cys
        465                 470                 475

act ggg gat gaa gat tac acc tat ttg att ctg cgg gtc att gga aat      1610
Thr Gly Asp Glu Asp Tyr Thr Tyr Leu Ile Leu Arg Val Ile Gly Asn
    480                 485                 490

atg ggc caa acc atg gag cag tta act cca gaa ctc aag tct tca atc      1658
Met Gly Gln Thr Met Glu Gln Leu Thr Pro Glu Leu Lys Ser Ser Ile
495                 500                 505                 510

ctc aaa tgt gtc caa agt aca aag cca tca ctg atg atc cag aaa gct      1706
Leu Lys Cys Val Gln Ser Thr Lys Pro Ser Leu Met Ile Gln Lys Ala
                515                 520                 525

gcc atc cag gct ctg cgg aaa atg gag cct aaa gac aag gac cag gag      1754
Ala Ile Gln Ala Leu Arg Lys Met Glu Pro Lys Asp Lys Asp Gln Glu
            530                 535                 540

gtt ctt ctt cag act ttc ctt gat gat gct tct ccg gga gat aag cga      1802
Val Leu Leu Gln Thr Phe Leu Asp Asp Ala Ser Pro Gly Asp Lys Arg
```

```
          545                     550                     555

ctg gct gcc tat ctt atg ttg atg agg agt cct tca cag gca gat att      1850
Leu Ala Ala Tyr Leu Met Leu Met Arg Ser Pro Ser Gln Ala Asp Ile
    560                     565                     570

aac aaa att gtc caa att cta cca tgg gaa cag aat gag caa gtg aag      1898
Asn Lys Ile Val Gln Ile Leu Pro Trp Glu Gln Asn Glu Gln Val Lys
575                     580                     585                 590

aac ttt gtg gct tcc cat att gcc aat atc ttg aac tca gaa gaa ttg      1946
Asn Phe Val Ala Ser His Ile Ala Asn Ile Leu Asn Ser Glu Glu Leu
                595                     600                     605

gat atc caa gat ctg aaa aag tta gtg aaa gaa gct ctg aaa gaa tct      1994
Asp Ile Gln Asp Leu Lys Lys Leu Val Lys Glu Ala Leu Lys Glu Ser
            610                     615                     620

caa ctt cca act gtc atg gac ttc aga aaa ttc tct cgg aac tat caa      2042
Gln Leu Pro Thr Val Met Asp Phe Arg Lys Phe Ser Arg Asn Tyr Gln
        625                     630                     635

ctc tac aaa tct gtt tct ctt cca tca ctt gac cca gcc tca gcc aaa      2090
Leu Tyr Lys Ser Val Ser Leu Pro Ser Leu Asp Pro Ala Ser Ala Lys
    640                     645                     650

ata gaa ggg aat ctt ata ttt gat cca aat aac tac ctt cct aaa gaa      2138
Ile Glu Gly Asn Leu Ile Phe Asp Pro Asn Asn Tyr Leu Pro Lys Glu
655                     660                     665                 670

agc atg ctg aaa act acc ctc act gcc ttt gga ttt gct tca gct gac      2186
Ser Met Leu Lys Thr Thr Leu Thr Ala Phe Gly Phe Ala Ser Ala Asp
                675                     680                     685

ctc atc gag att ggc ttg gaa gga aaa ggc ttt gag cca aca ttg gaa      2234
Leu Ile Glu Ile Gly Leu Glu Gly Lys Gly Phe Glu Pro Thr Leu Glu
            690                     695                     700

gct ctt ttt ggg aag caa gga ttt ttc cca gac agt gtc aac aaa gct      2282
Ala Leu Phe Gly Lys Gln Gly Phe Phe Pro Asp Ser Val Asn Lys Ala
        705                     710                     715

ttg tac tgg gtt aat ggt caa gtt cct gat ggt gtc tct aag gtc tta      2330
Leu Tyr Trp Val Asn Gly Gln Val Pro Asp Gly Val Ser Lys Val Leu
    720                     725                     730

gtg gac cac ttt ggc tat acc aaa gat gat aaa cat gag cag gat atg      2378
Val Asp His Phe Gly Tyr Thr Lys Asp Asp Lys His Glu Gln Asp Met
735                     740                     745                 750

gta aat gga ata atg ctc agt gtt gag aag ctg att aaa gat ttg aaa      2426
Val Asn Gly Ile Met Leu Ser Val Glu Lys Leu Ile Lys Asp Leu Lys
                755                     760                     765

tcc aaa gaa gtc ccg gaa gcc aga gcc tac ctc cgc atc ttg gga gag      2474
Ser Lys Glu Val Pro Glu Ala Arg Ala Tyr Leu Arg Ile Leu Gly Glu
            770                     775                     780
```

```
gag ctt ggt ttt gcc agt ctc cat gac ctc cag ctc ctg gga aag ctg      2522
Glu Leu Gly Phe Ala Ser Leu His Asp Leu Gln Leu Leu Gly Lys Leu
        785                 790                 795

ctt ctg atg ggt gcc cgc act ctg cag ggg atc ccc cag atg att gga      2570
Leu Leu Met Gly Ala Arg Thr Leu Gln Gly Ile Pro Gln Met Ile Gly
    800                 805                 810

gag gtc atc agg aag ggc tca aag aat gac ttt ttt ctt cac tac atc      2618
Glu Val Ile Arg Lys Gly Ser Lys Asn Asp Phe Phe Leu His Tyr Ile
815                 820                 825                 830

ttc atg gag aat gcc ttt gaa ctc ccc act gga gct gga tta cag ttg      2666
Phe Met Glu Asn Ala Phe Glu Leu Pro Thr Gly Ala Gly Leu Gln Leu
                835                 840                 845

caa ata tct tca tct gga gtc att gct ccc gga gcc aag gct gga gta      2714
Gln Ile Ser Ser Ser Gly Val Ile Ala Pro Gly Ala Lys Ala Gly Val
            850                 855                 860

aaa ctg gaa gta gcc aac atg cag gct gaa ctg gtg gca aaa ccc tcc      2762
Lys Leu Glu Val Ala Asn Met Gln Ala Glu Leu Val Ala Lys Pro Ser
        865                 870                 875

gtg tct gtg gag ttt gtg aca aat atg ggc atc atc att ccg gac ttc      2810
Val Ser Val Glu Phe Val Thr Asn Met Gly Ile Ile Ile Pro Asp Phe
    880                 885                 890

gct agg agt ggg gtc cag atg aac acc aac ttc ttc cac gag tcg ggt      2858
Ala Arg Ser Gly Val Gln Met Asn Thr Asn Phe Phe His Glu Ser Gly
895                 900                 905                 910

ctg gag gct cat gtt gcc cta aaa gct ggg aag ctg aag ttt atc att      2906
Leu Glu Ala His Val Ala Leu Lys Ala Gly Lys Leu Lys Phe Ile Ile
                915                 920                 925

cct tcc cca aag aga cca gtc aag ctg ctc agt gga ggc aac aca tta      2954
Pro Ser Pro Lys Arg Pro Val Lys Leu Leu Ser Gly Gly Asn Thr Leu
            930                 935                 940

cat ttg gtc tct acc acc aaa acg gag gtg atc cca cct ctc att gag      3002
His Leu Val Ser Thr Thr Lys Thr Glu Val Ile Pro Pro Leu Ile Glu
        945                 950                 955

aac agg cag tcc tgg tca gtt tgc aag caa gtc ttt cct ggc ctg aat      3050
Asn Arg Gln Ser Trp Ser Val Cys Lys Gln Val Phe Pro Gly Leu Asn
    960                 965                 970

tac tgc acc tca ggc gct tac tcc aac gcc agc tcc aca gac tcc gcc      3098
Tyr Cys Thr Ser Gly Ala Tyr Ser Asn Ala Ser Ser Thr Asp Ser Ala
975                 980                 985                 990

tcc tac tat ccg ctg acc ggg gac acc aga tta gag ctg gaa ctg agg      3146
Ser Tyr Tyr Pro Leu Thr Gly Asp Thr Arg Leu Glu Leu Glu Leu Arg
                995                 1000                1005

cct aca gga gag att gag cag tat tct gtc agc gca acc tat gag ctc      3194
Pro Thr Gly Glu Ile Glu Gln Tyr Ser Val Ser Ala Thr Tyr Glu Leu
```

```
            1010                1015                1020

cag aga gag gac aga gcc ttg gtg gat acc ctg aag ttt gta act caa      3242
Gln Arg Glu Asp Arg Ala Leu Val Asp Thr Leu Lys Phe Val Thr Gln
        1025                1030                1035

gca gaa ggt gcg aag cag act gag gct acc atg aca ttc aaa tat aat      3290
Ala Glu Gly Ala Lys Gln Thr Glu Ala Thr Met Thr Phe Lys Tyr Asn
    1040                1045                1050

cgg cag agt atg acc ttg tcc agt gaa gtc caa att ccg gat ttt gat      3338
Arg Gln Ser Met Thr Leu Ser Ser Glu Val Gln Ile Pro Asp Phe Asp
1055                1060                1065                1070

gtt gac ctc gga aca atc ctc aga gtt aat gat gaa tct act gag ggc      3386
Val Asp Leu Gly Thr Ile Leu Arg Val Asn Asp Glu Ser Thr Glu Gly
                1075                1080                1085

aaa acg tct tac aga ctc acc ctg gac att cag aac aag aaa att act      3434
Lys Thr Ser Tyr Arg Leu Thr Leu Asp Ile Gln Asn Lys Lys Ile Thr
            1090                1095                1100

gag gtc gcc ctc atg ggc cac cta agt tgt gac aca aag gaa gaa aga      3482
Glu Val Ala Leu Met Gly His Leu Ser Cys Asp Thr Lys Glu Glu Arg
        1105                1110                1115

aaa atc aag ggt gtt att tcc ata ccc cgt ttg caa gca gaa gcc aga      3530
Lys Ile Lys Gly Val Ile Ser Ile Pro Arg Leu Gln Ala Glu Ala Arg
    1120                1125                1130

agt gag atc ctc gcc cac tgg tcg cct gcc aaa ctg ctt ctc caa atg      3578
Ser Glu Ile Leu Ala His Trp Ser Pro Ala Lys Leu Leu Leu Gln Met
1135                1140                1145                1150

gac tca tct gct aca gct tat ggc tcc aca gtt tcc aag agg gtg gca      3626
Asp Ser Ser Ala Thr Ala Tyr Gly Ser Thr Val Ser Lys Arg Val Ala
                1155                1160                1165

tgg cat tat gat gaa gag aag att gaa ttt gaa tgg aac aca ggc acc      3674
Trp His Tyr Asp Glu Glu Lys Ile Glu Phe Glu Trp Asn Thr Gly Thr
            1170                1175                1180

aat gta gat acc aaa aaa atg act tcc aat ttc cct gtg gat ctc tcc      3722
Asn Val Asp Thr Lys Lys Met Thr Ser Asn Phe Pro Val Asp Leu Ser
        1185                1190                1195

gat tat cct aag agc ttg cat atg tat gct aat aga ctc ctg gat cac      3770
Asp Tyr Pro Lys Ser Leu His Met Tyr Ala Asn Arg Leu Leu Asp His
    1200                1205                1210

aga gtc cct gaa aca gac atg act ttc cgg cac gtg ggt tcc aaa tta      3818
Arg Val Pro Glu Thr Asp Met Thr Phe Arg His Val Gly Ser Lys Leu
1215                1220                1225                1230

ata gtt gca atg agc tca tgg ctt cag aag gca tct ggg agt ctt cct      3866
Ile Val Ala Met Ser Ser Trp Leu Gln Lys Ala Ser Gly Ser Leu Pro
                1235                1240                1245
```

```
tat acc cag act ttg caa gac cac ctc aat agc ctg aag gag ttc aac      3914
Tyr Thr Gln Thr Leu Gln Asp His Leu Asn Ser Leu Lys Glu Phe Asn
            1250                1255                1260

ctc cag aac atg gga ttg cca gac ttc cac atc cca gaa aac ctc ttc      3962
Leu Gln Asn Met Gly Leu Pro Asp Phe His Ile Pro Glu Asn Leu Phe
        1265                1270                1275

tta aaa agc gat ggc cgg gtc aaa tat acc ttg aac aag aac agt ttg      4010
Leu Lys Ser Asp Gly Arg Val Lys Tyr Thr Leu Asn Lys Asn Ser Leu
    1280                1285                1290

aaa att gag att cct ttg cct ttt ggt ggc aaa tcc tcc aga gat cta      4058
Lys Ile Glu Ile Pro Leu Pro Phe Gly Gly Lys Ser Ser Arg Asp Leu
1295                1300                1305                1310

aag atg tta gag act gtt agg aca cca gcc ctc cac ttc aag tct gtg      4106
Lys Met Leu Glu Thr Val Arg Thr Pro Ala Leu His Phe Lys Ser Val
                1315                1320                1325

gga ttc cat ctg cca tct cga gag ttc caa gtc cct act ttt acc att      4154
Gly Phe His Leu Pro Ser Arg Glu Phe Gln Val Pro Thr Phe Thr Ile
            1330                1335                1340

ccc aag ttg tat caa ctg caa gtg cct ctc ctg ggt gtt cta gac ctc      4202
Pro Lys Leu Tyr Gln Leu Gln Val Pro Leu Leu Gly Val Leu Asp Leu
        1345                1350                1355

tcc acg aat gtc tac agc aac ttg tac aac tgg tcc gcc tcc tac agt      4250
Ser Thr Asn Val Tyr Ser Asn Leu Tyr Asn Trp Ser Ala Ser Tyr Ser
    1360                1365                1370

ggt ggc aac acc agc aca gac cat ttc agc ctt cgg gct cgt tac cac      4298
Gly Gly Asn Thr Ser Thr Asp His Phe Ser Leu Arg Ala Arg Tyr His
1375                1380                1385                1390

atg aag gct gac tct gtg gtt gac ctg ctt tcc tac aat gtg caa gga      4346
Met Lys Ala Asp Ser Val Val Asp Leu Leu Ser Tyr Asn Val Gln Gly
                1395                1400                1405

tct gga gaa aca aca tat gac cac aag aat acg ttc aca cta tca tgt      4394
Ser Gly Glu Thr Thr Tyr Asp His Lys Asn Thr Phe Thr Leu Ser Cys
            1410                1415                1420

gat ggg tct cta cgc cac aaa ttt cta gat tcg aat atc aaa ttc agt      4442
Asp Gly Ser Leu Arg His Lys Phe Leu Asp Ser Asn Ile Lys Phe Ser
        1425                1430                1435

cat gta gaa aaa ctt gga aac aac cca gtc tca aaa ggt tta cta ata      4490
His Val Glu Lys Leu Gly Asn Asn Pro Val Ser Lys Gly Leu Leu Ile
    1440                1445                1450

ttc gat gca tct agt tcc tgg gga cca cag atg tct gct tca gtt cat      4538
Phe Asp Ala Ser Ser Ser Trp Gly Pro Gln Met Ser Ala Ser Val His
1455                1460                1465                1470

ttg gac tcc aaa aag aaa cag cat ttg ttt gtc aaa gaa gtc aag att      4586
Leu Asp Ser Lys Lys Lys Gln His Leu Phe Val Lys Glu Val Lys Ile
```

```
                1475                1480                1485

gat ggg cag ttc aga gtc tct tcg ttc tat gct aaa ggc aca tat ggc      4634
Asp Gly Gln Phe Arg Val Ser Ser Phe Tyr Ala Lys Gly Thr Tyr Gly
            1490                1495                1500

ctg tct tgt cag agg gat cct aac act ggc cgg ctc aat gga gag tcc      4682
Leu Ser Cys Gln Arg Asp Pro Asn Thr Gly Arg Leu Asn Gly Glu Ser
        1505                1510                1515

aac ctg agg ttt aac tcc tcc tac ctc caa ggc acc aac cag ata aca      4730
Asn Leu Arg Phe Asn Ser Ser Tyr Leu Gln Gly Thr Asn Gln Ile Thr
    1520                1525                1530

gga aga tat gaa gat gga acc ctc tcc ctc acc tcc acc tct gat ctg      4778
Gly Arg Tyr Glu Asp Gly Thr Leu Ser Leu Thr Ser Thr Ser Asp Leu
1535                1540                1545                1550

caa agt ggc atc att aaa aat act gct tcc cta aag tat gag aac tac      4826
Gln Ser Gly Ile Ile Lys Asn Thr Ala Ser Leu Lys Tyr Glu Asn Tyr
                1555                1560                1565

gag ctg act tta aaa tct gac acc aat ggg aag tat aag aac ttt gcc      4874
Glu Leu Thr Leu Lys Ser Asp Thr Asn Gly Lys Tyr Lys Asn Phe Ala
            1570                1575                1580

act tct aac aag atg gat atg acc ttc tct aag caa aat gca ctg ctg      4922
Thr Ser Asn Lys Met Asp Met Thr Phe Ser Lys Gln Asn Ala Leu Leu
        1585                1590                1595

cgt tct gaa tat cag gct gat tac gag tca ttg agg ttc ttc agc ctg      4970
Arg Ser Glu Tyr Gln Ala Asp Tyr Glu Ser Leu Arg Phe Phe Ser Leu
    1600                1605                1610

ctt tct gga tca cta aat tcc cat ggt ctt gag tta aat gct gac atc      5018
Leu Ser Gly Ser Leu Asn Ser His Gly Leu Glu Leu Asn Ala Asp Ile
1615                1620                1625                1630

tta ggc act gac aaa att aat agt ggt gct cac aag gcg aca cta agg      5066
Leu Gly Thr Asp Lys Ile Asn Ser Gly Ala His Lys Ala Thr Leu Arg
                1635                1640                1645

att ggc caa gat gga ata tct acc agt gca acg acc aac ttg aag tgt      5114
Ile Gly Gln Asp Gly Ile Ser Thr Ser Ala Thr Thr Asn Leu Lys Cys
            1650                1655                1660

agt ctc ctg gtg ctg gag aat gag ctg aat gca gag ctt ggc ctc tct      5162
Ser Leu Leu Val Leu Glu Asn Glu Leu Asn Ala Glu Leu Gly Leu Ser
        1665                1670                1675

ggg gca tct atg aaa tta aca aca aat ggc cgc ttc agg gaa cac aat      5210
Gly Ala Ser Met Lys Leu Thr Thr Asn Gly Arg Phe Arg Glu His Asn
    1680                1685                1690

gca aaa ttc agt ctg gat ggg aaa gcc gcc ctc aca gag cta tca ctg      5258
Ala Lys Phe Ser Leu Asp Gly Lys Ala Ala Leu Thr Glu Leu Ser Leu
1695                1700                1705                1710
```

```
gga agt gct tat cag gcc atg att ctg ggt gtc gac agc aaa aac att      5306
Gly Ser Ala Tyr Gln Ala Met Ile Leu Gly Val Asp Ser Lys Asn Ile
                1715                1720                1725

ttc aac ttc aag gtc agt caa gaa gga ctt aag ctc tca aat gac atg      5354
Phe Asn Phe Lys Val Ser Gln Glu Gly Leu Lys Leu Ser Asn Asp Met
            1730                1735                1740

atg ggc tca tat gct gaa atg aaa ttt gac cac aca aac agt ctg aac      5402
Met Gly Ser Tyr Ala Glu Met Lys Phe Asp His Thr Asn Ser Leu Asn
        1745                1750                1755

att gca ggc tta tca ctg gac ttc tct tca aaa ctt gac aac att tac      5450
Ile Ala Gly Leu Ser Leu Asp Phe Ser Ser Lys Leu Asp Asn Ile Tyr
    1760                1765                1770

agc tct gac aag ttt tat aag caa act gtt aat tta cag cta cag ccc      5498
Ser Ser Asp Lys Phe Tyr Lys Gln Thr Val Asn Leu Gln Leu Gln Pro
1775                1780                1785                1790

tat tct ctg gta act act tta aac agt gac ctg aaa tac aat gct ctg      5546
Tyr Ser Leu Val Thr Thr Leu Asn Ser Asp Leu Lys Tyr Asn Ala Leu
                1795                1800                1805

gat ctc acc aac aat ggg aaa cta cgg cta gaa ccc ctg aag ctg cat      5594
Asp Leu Thr Asn Asn Gly Lys Leu Arg Leu Glu Pro Leu Lys Leu His
            1810                1815                1820

gtg gct ggt aac cta aaa gga gcc tac caa aat aat gaa ata aaa cac      5642
Val Ala Gly Asn Leu Lys Gly Ala Tyr Gln Asn Asn Glu Ile Lys His
        1825                1830                1835

atc tat gcc atc tct tct gct gcc tta tca gca agc tat aaa gca gac      5690
Ile Tyr Ala Ile Ser Ser Ala Ala Leu Ser Ala Ser Tyr Lys Ala Asp
    1840                1845                1850

act gtt gct aag gtt cag ggt gtg gag ttt agc cat cgg ctc aac aca      5738
Thr Val Ala Lys Val Gln Gly Val Glu Phe Ser His Arg Leu Asn Thr
1855                1860                1865                1870

gac atc gct ggg ctg gct tca gcc att gac atg agc aca aac tat aat      5786
Asp Ile Ala Gly Leu Ala Ser Ala Ile Asp Met Ser Thr Asn Tyr Asn
                1875                1880                1885

tca gac tca ctg cat ttc agc aat gtc ttc cgt tct gta atg gcc ccg      5834
Ser Asp Ser Leu His Phe Ser Asn Val Phe Arg Ser Val Met Ala Pro
            1890                1895                1900

ttt acc atg acc atc gat gca cat aca aat ggc aat ggg aaa ctc gct      5882
Phe Thr Met Thr Ile Asp Ala His Thr Asn Gly Asn Gly Lys Leu Ala
        1905                1910                1915

ctc tgg gga gaa cat act ggg cag ctg tat agc aaa ttc ctg ttg aaa      5930
Leu Trp Gly Glu His Thr Gly Gln Leu Tyr Ser Lys Phe Leu Leu Lys
    1920                1925                1930

gca gaa cct ctg gca ttt act ttc tct cat gat tac aaa ggc tcc aca      5978
Ala Glu Pro Leu Ala Phe Thr Phe Ser His Asp Tyr Lys Gly Ser Thr
```

```
1935                1940                1945                1950

agt cat cat ctc gtg tct agg aaa agc atc agt gca gct ctt gaa cac      6026
Ser His His Leu Val Ser Arg Lys Ser Ile Ser Ala Ala Leu Glu His
                1955                1960                1965

aaa gtc agt gcc ctg ctt act cca gct gag cag aca ggc acc tgg aaa      6074
Lys Val Ser Ala Leu Leu Thr Pro Ala Glu Gln Thr Gly Thr Trp Lys
            1970                1975                1980

ctc aag acc caa ttt aac aac aat gaa tac agc cag gac ttg gat gct      6122
Leu Lys Thr Gln Phe Asn Asn Asn Glu Tyr Ser Gln Asp Leu Asp Ala
        1985                1990                1995

tac aac act aaa gat aaa att ggc gtg gag ctt act gga cga act ctg      6170
Tyr Asn Thr Lys Asp Lys Ile Gly Val Glu Leu Thr Gly Arg Thr Leu
    2000                2005                2010

gct gac cta act cta cta gac tcc cca att aaa gtg cca ctt tta ctc      6218
Ala Asp Leu Thr Leu Leu Asp Ser Pro Ile Lys Val Pro Leu Leu Leu
2015                2020                2025                2030

agt gag ccc atc aat atc att gat gct tta gag atg aga gat gcc gtt      6266
Ser Glu Pro Ile Asn Ile Ile Asp Ala Leu Glu Met Arg Asp Ala Val
                2035                2040                2045

gag aag ccc caa gaa ttt aca att gtt gct ttt gta aag tat gat aaa      6314
Glu Lys Pro Gln Glu Phe Thr Ile Val Ala Phe Val Lys Tyr Asp Lys
            2050                2055                2060

aac caa gat gtt cac tcc att aac ctc cca ttt ttt gag acc ttg caa      6362
Asn Gln Asp Val His Ser Ile Asn Leu Pro Phe Phe Glu Thr Leu Gln
        2065                2070                2075

gaa tat ttt gag agg aat cga caa acc att ata gtt gta gtg gaa aac      6410
Glu Tyr Phe Glu Arg Asn Arg Gln Thr Ile Ile Val Val Val Glu Asn
    2080                2085                2090

gta cag aga aac ctg aag cac atc aat att gat caa ttt gta aga aaa      6458
Val Gln Arg Asn Leu Lys His Ile Asn Ile Asp Gln Phe Val Arg Lys
2095                2100                2105                2110

tac aga gca gcc ctg gga aaa ctc cca cag caa gct aat gat tat ctg      6506
Tyr Arg Ala Ala Leu Gly Lys Leu Pro Gln Gln Ala Asn Asp Tyr Leu
                2115                2120                2125

aat tca ttc aat tgg gag aga caa gtt tca cat gcc aag gag aaa ctg      6554
Asn Ser Phe Asn Trp Glu Arg Gln Val Ser His Ala Lys Glu Lys Leu
            2130                2135                2140

act gct ctc aca aaa aag tat aga att aca gaa aat gat ata caa att      6602
Thr Ala Leu Thr Lys Lys Tyr Arg Ile Thr Glu Asn Asp Ile Gln Ile
        2145                2150                2155

gca tta gat gat gcc aaa atc aac ttt aat gaa aaa cta tct caa ctg      6650
Ala Leu Asp Asp Ala Lys Ile Asn Phe Asn Glu Lys Leu Ser Gln Leu
    2160                2165                2170
```

```
cag aca tat atg ata caa ttt gat cag tat att aaa gat agt tat gat     6698
Gln Thr Tyr Met Ile Gln Phe Asp Gln Tyr Ile Lys Asp Ser Tyr Asp
2175                2180                2185                2190

tta cat gat ttg aaa ata gct att gct aat att att gat gaa atc att     6746
Leu His Asp Leu Lys Ile Ala Ile Ala Asn Ile Ile Asp Glu Ile Ile
                2195                2200                2205

gaa aaa tta aaa agt ctt gat gag cac tat cat atc cgt gta aat tta     6794
Glu Lys Leu Lys Ser Leu Asp Glu His Tyr His Ile Arg Val Asn Leu
            2210                2215                2220

gta aaa aca atc cat gat cta cat ttg ttt att gaa aat att gat ttt     6842
Val Lys Thr Ile His Asp Leu His Leu Phe Ile Glu Asn Ile Asp Phe
        2225                2230                2235

aac aaa agt gga agt agt act gca tcc tgg att caa aat gtg gat act     6890
Asn Lys Ser Gly Ser Ser Thr Ala Ser Trp Ile Gln Asn Val Asp Thr
    2240                2245                2250

aag tac caa atc aga atc cag ata caa gaa aaa ctg cag cag ctt aag     6938
Lys Tyr Gln Ile Arg Ile Gln Ile Gln Glu Lys Leu Gln Gln Leu Lys
2255                2260                2265                2270

aga cac ata cag aat ata gac atc cag cac cta gct gga aag tta aaa     6986
Arg His Ile Gln Asn Ile Asp Ile Gln His Leu Ala Gly Lys Leu Lys
                2275                2280                2285

caa cac att gag gct att gat gtt aga gtg ctt tta gat caa ttg gga     7034
Gln His Ile Glu Ala Ile Asp Val Arg Val Leu Leu Asp Gln Leu Gly
            2290                2295                2300

act aca att tca ttt gaa aga ata aat gat gtt ctt gag cat gtc aaa     7082
Thr Thr Ile Ser Phe Glu Arg Ile Asn Asp Val Leu Glu His Val Lys
        2305                2310                2315

cac ttt gtt ata aat ctt att ggg gat ttt gaa gta gct gag aaa atc     7130
His Phe Val Ile Asn Leu Ile Gly Asp Phe Glu Val Ala Glu Lys Ile
    2320                2325                2330

aat gcc ttc aga gcc aaa gtc cat gag tta atc gag agg tat gaa gta     7178
Asn Ala Phe Arg Ala Lys Val His Glu Leu Ile Glu Arg Tyr Glu Val
2335                2340                2345                2350

gac caa caa atc cag gtt tta atg gat aaa tta gta gag ttg acc cac     7226
Asp Gln Gln Ile Gln Val Leu Met Asp Lys Leu Val Glu Leu Thr His
                2355                2360                2365

caa tac aag ttg aag gag act att cag aag cta agc aat gtc cta caa     7274
Gln Tyr Lys Leu Lys Glu Thr Ile Gln Lys Leu Ser Asn Val Leu Gln
            2370                2375                2380

caa gtt aag ata aaa gat tac ttt gag aaa ttg gtt gga ttt att gat     7322
Gln Val Lys Ile Lys Asp Tyr Phe Glu Lys Leu Val Gly Phe Ile Asp
        2385                2390                2395

gat gct gtg aag aag ctt aat gaa tta tct ttt aaa aca ttc att gaa     7370
Asp Ala Val Lys Lys Leu Asn Glu Leu Ser Phe Lys Thr Phe Ile Glu
```

```
    2400                    2405                    2410

gat gtt aac aaa ttc ctt gac atg ttg ata aag aaa tta aag tca ttt      7418
Asp Val Asn Lys Phe Leu Asp Met Leu Ile Lys Lys Leu Lys Ser Phe
2415                2420                2425                2430

gat tac cac cag ttt gta gat gaa acc aat gac aaa atc cgt gag gtg      7466
Asp Tyr His Gln Phe Val Asp Glu Thr Asn Asp Lys Ile Arg Glu Val
                2435                2440                2445

act cag aga ctc aat ggt gaa att cag gct ctg gaa cta cca caa aaa      7514
Thr Gln Arg Leu Asn Gly Glu Ile Gln Ala Leu Glu Leu Pro Gln Lys
            2450                2455                2460

gct gaa gca tta aaa ctg ttt tta gag gaa acc aag gcc aca gtt gca      7562
Ala Glu Ala Leu Lys Leu Phe Leu Glu Glu Thr Lys Ala Thr Val Ala
        2465                2470                2475

gtg tat ctg gaa agc cta cag gac acc aaa ata acc tta atc atc aat      7610
Val Tyr Leu Glu Ser Leu Gln Asp Thr Lys Ile Thr Leu Ile Ile Asn
    2480                2485                2490

tgg tta cag gag gct tta agt tca gca tct ttg gct cac atg aag gcc      7658
Trp Leu Gln Glu Ala Leu Ser Ser Ala Ser Leu Ala His Met Lys Ala
2495                2500                2505                2510

aaa ttc cga gag act cta gaa gat aca cga gac cga atg tat caa atg      7706
Lys Phe Arg Glu Thr Leu Glu Asp Thr Arg Asp Arg Met Tyr Gln Met
                2515                2520                2525

gac att cag cag gaa ctt caa cga tac ctg tct ctg gta ggc cag gtt      7754
Asp Ile Gln Gln Glu Leu Gln Arg Tyr Leu Ser Leu Val Gly Gln Val
            2530                2535                2540

tat agc aca ctt gtc acc tac att tct gat tgg tgg act ctt gct gct      7802
Tyr Ser Thr Leu Val Thr Tyr Ile Ser Asp Trp Trp Thr Leu Ala Ala
        2545                2550                2555

aag aac ctt act gac ttt gca gag caa tat tct atc caa gat tgg gct      7850
Lys Asn Leu Thr Asp Phe Ala Glu Gln Tyr Ser Ile Gln Asp Trp Ala
    2560                2565                2570

aaa cgt atg aaa gca ttg gta gag caa ggg ttc act gtt cct gaa atc      7898
Lys Arg Met Lys Ala Leu Val Glu Gln Gly Phe Thr Val Pro Glu Ile
2575                2580                2585                2590

aag acc atc ctt ggg acc atg cct gcc ttt gaa gtc agt ctt cag gct      7946
Lys Thr Ile Leu Gly Thr Met Pro Ala Phe Glu Val Ser Leu Gln Ala
                2595                2600                2605

ctt cag aaa gct acc ttc cag aca cct gat ttt ata gtc ccc cta aca      7994
Leu Gln Lys Ala Thr Phe Gln Thr Pro Asp Phe Ile Val Pro Leu Thr
            2610                2615                2620

gat ttg agg att cca tca gtt cag ata aac ttc aaa gac tta aaa aat      8042
Asp Leu Arg Ile Pro Ser Val Gln Ile Asn Phe Lys Asp Leu Lys Asn
        2625                2630                2635
```

```
ata aaa atc cca tcc agg ttt tcc aca cca gaa ttt acc atc ctt aac      8090
Ile Lys Ile Pro Ser Arg Phe Ser Thr Pro Glu Phe Thr Ile Leu Asn
    2640                2645                2650

acc ttc cac att cct tcc ttt aca att gac ttt gtc gaa atg aaa gta      8138
Thr Phe His Ile Pro Ser Phe Thr Ile Asp Phe Val Glu Met Lys Val
2655                2660                2665                2670

aag atc atc aga acc att gac cag atg cag aac agt gag ctg cag tgg      8186
Lys Ile Ile Arg Thr Ile Asp Gln Met Gln Asn Ser Glu Leu Gln Trp
                2675                2680                2685

ccc gtt cca gat ata tat ctc agg gat ctg aag gtg gag gac att cct      8234
Pro Val Pro Asp Ile Tyr Leu Arg Asp Leu Lys Val Glu Asp Ile Pro
            2690                2695                2700

cta gcg aga atc acc ctg cca gac ttc cgt tta cca gaa atc gca att      8282
Leu Ala Arg Ile Thr Leu Pro Asp Phe Arg Leu Pro Glu Ile Ala Ile
        2705                2710                2715

cca gaa ttc ata atc cca act ctc aac ctt aat gat ttt caa gtt cct      8330
Pro Glu Phe Ile Ile Pro Thr Leu Asn Leu Asn Asp Phe Gln Val Pro
    2720                2725                2730

gac ctt cac ata cca gaa ttc cag ctt ccc cac atc tca cac aca att      8378
Asp Leu His Ile Pro Glu Phe Gln Leu Pro His Ile Ser His Thr Ile
2735                2740                2745                2750

gaa gta cct act ttt ggc aag cta tac agt att ctg aaa atc caa tct      8426
Glu Val Pro Thr Phe Gly Lys Leu Tyr Ser Ile Leu Lys Ile Gln Ser
                2755                2760                2765

cct ctt ttc aca tta gat gca aat gct gac ata ggg aat gga acc acc      8474
Pro Leu Phe Thr Leu Asp Ala Asn Ala Asp Ile Gly Asn Gly Thr Thr
            2770                2775                2780

tca gca aac gaa gca ggt atc gca gct tcc atc act gcc aaa gga gag      8522
Ser Ala Asn Glu Ala Gly Ile Ala Ala Ser Ile Thr Ala Lys Gly Glu
        2785                2790                2795

tcc aaa tta gaa gtt ctc aat ttt gat ttt caa gca aat gca caa ctc      8570
Ser Lys Leu Glu Val Leu Asn Phe Asp Phe Gln Ala Asn Ala Gln Leu
    2800                2805                2810

tca aac cct aag att aat ccg ctg gct ctg aag gag tca gtg aag ttc      8618
Ser Asn Pro Lys Ile Asn Pro Leu Ala Leu Lys Glu Ser Val Lys Phe
2815                2820                2825                2830

tcc agc aag tac ctg aga acg gag cat ggg agt gaa atg ctg ttt ttt      8666
Ser Ser Lys Tyr Leu Arg Thr Glu His Gly Ser Glu Met Leu Phe Phe
                2835                2840                2845

gga aat gct att gag gga aaa tca aac aca gtg gca agt tta cac aca      8714
Gly Asn Ala Ile Glu Gly Lys Ser Asn Thr Val Ala Ser Leu His Thr
            2850                2855                2860

gaa aaa aat aca ctg gag ctt agt aat gga gtg att gtc aag ata aac      8762
Glu Lys Asn Thr Leu Glu Leu Ser Asn Gly Val Ile Val Lys Ile Asn
```

```
        2865                    2870                    2875

aat cag ctt acc ctg gat agc aac act aaa tac ttc cac aaa ttg aac      8810
Asn Gln Leu Thr Leu Asp Ser Asn Thr Lys Tyr Phe His Lys Leu Asn
    2880                    2885                    2890

atc ccc aaa ctg gac ttc tct agt cag gct gac ctg cgc aac gag atc      8858
Ile Pro Lys Leu Asp Phe Ser Ser Gln Ala Asp Leu Arg Asn Glu Ile
2895                    2900                    2905                    2910

aag aca ctg ttg aaa gct ggc cac ata gca tgg act tct tct gga aaa      8906
Lys Thr Leu Leu Lys Ala Gly His Ile Ala Trp Thr Ser Ser Gly Lys
                2915                    2920                    2925

ggg tca tgg aaa tgg gcc tgc ccc aga ttc tca gat gag gga aca cat      8954
Gly Ser Trp Lys Trp Ala Cys Pro Arg Phe Ser Asp Glu Gly Thr His
            2930                    2935                    2940

gaa tca caa att agt ttc acc ata gaa gga ccc ctc act tcc ttt gga      9002
Glu Ser Gln Ile Ser Phe Thr Ile Glu Gly Pro Leu Thr Ser Phe Gly
        2945                    2950                    2955

ctg tcc aat aag atc aat agc aaa cac cta aga gta aac caa aac ttg      9050
Leu Ser Asn Lys Ile Asn Ser Lys His Leu Arg Val Asn Gln Asn Leu
    2960                    2965                    2970

gtt tat gaa tct ggc tcc ctc aac ttt tct aaa ctt gaa att caa tca      9098
Val Tyr Glu Ser Gly Ser Leu Asn Phe Ser Lys Leu Glu Ile Gln Ser
2975                    2980                    2985                    2990

caa gtc gat tcc cag cat gtg ggc cac agt gtt cta act gct aaa ggc      9146
Gln Val Asp Ser Gln His Val Gly His Ser Val Leu Thr Ala Lys Gly
                2995                    3000                    3005

atg gca ctg ttt gga gaa ggg aag gca gag ttt act ggg agg cat gat      9194
Met Ala Leu Phe Gly Glu Gly Lys Ala Glu Phe Thr Gly Arg His Asp
            3010                    3015                    3020

gct cat tta aat gga aag gtt att gga act ttg aaa aat tct ctt ttc      9242
Ala His Leu Asn Gly Lys Val Ile Gly Thr Leu Lys Asn Ser Leu Phe
        3025                    3030                    3035

ttt tca gcc cag cca ttt gag atc acg gca tcc aca aac aat gaa ggg      9290
Phe Ser Ala Gln Pro Phe Glu Ile Thr Ala Ser Thr Asn Asn Glu Gly
    3040                    3045                    3050

aat ttg aaa gtt cgt ttt cca tta agg tta aca ggg aag ata gac ttc      9338
Asn Leu Lys Val Arg Phe Pro Leu Arg Leu Thr Gly Lys Ile Asp Phe
3055                    3060                    3065                    3070

ctg aat aac tat gca ctg ttt ctg agt ccc agt gcc cag caa gca agt      9386
Leu Asn Asn Tyr Ala Leu Phe Leu Ser Pro Ser Ala Gln Gln Ala Ser
                3075                    3080                    3085

tgg caa gta agt gct agg ttc aat cag tat aag tac aac caa aat ttc      9434
Trp Gln Val Ser Ala Arg Phe Asn Gln Tyr Lys Tyr Asn Gln Asn Phe
            3090                    3095                    3100
```

```
tct gct gga aac aac gag aac att atg gag gcc cat gta gga ata aat      9482
Ser Ala Gly Asn Asn Glu Asn Ile Met Glu Ala His Val Gly Ile Asn
        3105                    3110                    3115

gga gaa gca aat ctg gat ttc tta aac att cct tta aca att cct gaa      9530
Gly Glu Ala Asn Leu Asp Phe Leu Asn Ile Pro Leu Thr Ile Pro Glu
    3120                    3125                    3130

atg cgt cta cct tac aca ata atc aca act cct cca ctg aaa gat ttc      9578
Met Arg Leu Pro Tyr Thr Ile Ile Thr Thr Pro Pro Leu Lys Asp Phe
3135                    3140                    3145                3150

tct cta tgg gaa aaa aca ggc ttg aag gaa ttc ttg aaa acg aca aag      9626
Ser Leu Trp Glu Lys Thr Gly Leu Lys Glu Phe Leu Lys Thr Thr Lys
                3155                    3160                    3165

caa tca ttt gat tta agt gta aaa gct cag tat aag aaa aac aaa cac      9674
Gln Ser Phe Asp Leu Ser Val Lys Ala Gln Tyr Lys Lys Asn Lys His
            3170                    3175                    3180

agg cat tcc atc aca aat cct ttg gct gtg ctt tgt gag ttt atc agt      9722
Arg His Ser Ile Thr Asn Pro Leu Ala Val Leu Cys Glu Phe Ile Ser
        3185                    3190                    3195

cag agc atc aaa tcc ttt gac agg cat ttt gaa aaa aac aga aac aat      9770
Gln Ser Ile Lys Ser Phe Asp Arg His Phe Glu Lys Asn Arg Asn Asn
    3200                    3205                    3210

gca tta gat ttt gtc acc aaa tcc tat aat gaa aca aaa att aag ttt      9818
Ala Leu Asp Phe Val Thr Lys Ser Tyr Asn Glu Thr Lys Ile Lys Phe
3215                    3220                    3225                3230

gat aag tac aaa gct gaa aaa tct cac gac gag ctc ccc agg acc ttt      9866
Asp Lys Tyr Lys Ala Glu Lys Ser His Asp Glu Leu Pro Arg Thr Phe
                3235                    3240                    3245

caa att cct gga tac act gtt cca gtt gtc aat gtt gaa gtg tct cca      9914
Gln Ile Pro Gly Tyr Thr Val Pro Val Val Asn Val Glu Val Ser Pro
            3250                    3255                    3260

ttc acc ata gag atg tcg gca ttc ggc tat gtg ttc cca aaa gca gtc      9962
Phe Thr Ile Glu Met Ser Ala Phe Gly Tyr Val Phe Pro Lys Ala Val
        3265                    3270                    3275

agc atg cct agt ttc tcc atc cta ggt tct gac gtc cgt gtg cct tca     10010
Ser Met Pro Ser Phe Ser Ile Leu Gly Ser Asp Val Arg Val Pro Ser
    3280                    3285                    3290

tac aca tta atc ctg cca tca tta gag ctg cca gtc ctt cat gtc cct     10058
Tyr Thr Leu Ile Leu Pro Ser Leu Glu Leu Pro Val Leu His Val Pro
3295                    3300                    3305                3310

aga aat ctc aag ctt tct ctt cca cat ttc aag gaa ttg tgt acc ata     10106
Arg Asn Leu Lys Leu Ser Leu Pro His Phe Lys Glu Leu Cys Thr Ile
                3315                    3320                    3325

agc cat att ttt att cct gcc atg ggc aat att acc tat gat ttc tcc     10154
Ser His Ile Phe Ile Pro Ala Met Gly Asn Ile Thr Tyr Asp Phe Ser
```

```
                3330                3335                3340

ttt aaa tca agt gtc atc aca ctg aat acc aat gct gaa ctt ttt aac     10202
Phe Lys Ser Ser Val Ile Thr Leu Asn Thr Asn Ala Glu Leu Phe Asn
            3345                3350                3355

cag tca gat att gtt gct cat ctc ctt tct tca tct tca tct gtc att     10250
Gln Ser Asp Ile Val Ala His Leu Leu Ser Ser Ser Ser Ser Val Ile
        3360                3365                3370

gat gca ctg cag tac aaa tta gag ggc acc aca aga ttg aca aga aaa     10298
Asp Ala Leu Gln Tyr Lys Leu Glu Gly Thr Thr Arg Leu Thr Arg Lys
3375                3380                3385                3390

agg gga ttg aag tta gcc aca gct ctg tct ctg agc aac aaa ttt gtg     10346
Arg Gly Leu Lys Leu Ala Thr Ala Leu Ser Leu Ser Asn Lys Phe Val
                3395                3400                3405

gag ggt agt cat aac agt act gtg agc tta acc acg aaa aat atg gaa     10394
Glu Gly Ser His Asn Ser Thr Val Ser Leu Thr Thr Lys Asn Met Glu
            3410                3415                3420

gtg tca gtg gca aaa acc aca aaa gcc gaa att cca att ttg aga atg     10442
Val Ser Val Ala Lys Thr Thr Lys Ala Glu Ile Pro Ile Leu Arg Met
        3425                3430                3435

aat ttc aag caa gaa ctt aat gga aat acc aag tca aaa cct act gtc     10490
Asn Phe Lys Gln Glu Leu Asn Gly Asn Thr Lys Ser Lys Pro Thr Val
    3440                3445                3450

tct tcc tcc atg gaa ttt aag tat gat ttc aat tct tca atg ctg tac     10538
Ser Ser Ser Met Glu Phe Lys Tyr Asp Phe Asn Ser Ser Met Leu Tyr
3455                3460                3465                3470

tct acc gct aaa gga gca gtt gac cac aag ctt agc ttg gaa agc ctc     10586
Ser Thr Ala Lys Gly Ala Val Asp His Lys Leu Ser Leu Glu Ser Leu
                3475                3480                3485

acc tct tac ttt tcc att gag tca tct acc aaa gga gat gtc aag ggt     10634
Thr Ser Tyr Phe Ser Ile Glu Ser Ser Thr Lys Gly Asp Val Lys Gly
            3490                3495                3500

tcg gtt ctt tct cgg gaa tat tca gga act att gct agt gag gcc aac     10682
Ser Val Leu Ser Arg Glu Tyr Ser Gly Thr Ile Ala Ser Glu Ala Asn
        3505                3510                3515

act tac ttg aat tcc aag agc aca cgg tct tca gtg aag ctg cag ggc     10730
Thr Tyr Leu Asn Ser Lys Ser Thr Arg Ser Ser Val Lys Leu Gln Gly
    3520                3525                3530

act tcc aaa att gat gat atc tgg aac ctt gaa gta aaa gaa aat ttt     10778
Thr Ser Lys Ile Asp Asp Ile Trp Asn Leu Glu Val Lys Glu Asn Phe
3535                3540                3545                3550

gct gga gaa gcc aca ctc caa cgc ata tat tcc ctc tgg gag cac agt     10826
Ala Gly Glu Ala Thr Leu Gln Arg Ile Tyr Ser Leu Trp Glu His Ser
                3555                3560                3565
```

```
acg aaa aac cac tta cag cta gag ggc ctc ttt ttc acc aac gga gaa     10874
Thr Lys Asn His Leu Gln Leu Glu Gly Leu Phe Phe Thr Asn Gly Glu
            3570                3575                3580

cat aca agc aaa gcc acc ctg gaa ctc tct cca tgg caa atg tca gct     10922
His Thr Ser Lys Ala Thr Leu Glu Leu Ser Pro Trp Gln Met Ser Ala
        3585                3590                3595

ctt gtt cag gtc cat gca agt cag ccc agt tcc ttc cat gat ttc cct     10970
Leu Val Gln Val His Ala Ser Gln Pro Ser Ser Phe His Asp Phe Pro
    3600                3605                3610

gac ctt ggc cag gaa gtg gcc ctg aat gct aac act aag aac cag aag     11018
Asp Leu Gly Gln Glu Val Ala Leu Asn Ala Asn Thr Lys Asn Gln Lys
3615                3620                3625                3630

atc aga tgg aaa aat gaa gtc cgg att cat tct ggg tct ttc cag agc     11066
Ile Arg Trp Lys Asn Glu Val Arg Ile His Ser Gly Ser Phe Gln Ser
                3635                3640                3645

cag gtc gag ctt tcc aat gac caa gaa aag gca cac ctt gac att gca     11114
Gln Val Glu Leu Ser Asn Asp Gln Glu Lys Ala His Leu Asp Ile Ala
            3650                3655                3660

gga tcc tta gaa gga cac cta agg ttc ctc aaa aat atc atc cta cca     11162
Gly Ser Leu Glu Gly His Leu Arg Phe Leu Lys Asn Ile Ile Leu Pro
        3665                3670                3675

gtc tat gac aag agc tta tgg gat ttc cta aag ctg gat gta acc acc     11210
Val Tyr Asp Lys Ser Leu Trp Asp Phe Leu Lys Leu Asp Val Thr Thr
    3680                3685                3690

agc att ggt agg aga cag cat ctt cgt gtt tca act gcc ttt gtg tac     11258
Ser Ile Gly Arg Arg Gln His Leu Arg Val Ser Thr Ala Phe Val Tyr
3695                3700                3705                3710

acc aaa aac ccc aat ggc tat tca ttc tcc atc cct gta aaa gtt ttg     11306
Thr Lys Asn Pro Asn Gly Tyr Ser Phe Ser Ile Pro Val Lys Val Leu
                3715                3720                3725

gct gat aaa ttc att act cct ggg ctg aaa cta aat gat cta aat tca     11354
Ala Asp Lys Phe Ile Thr Pro Gly Leu Lys Leu Asn Asp Leu Asn Ser
            3730                3735                3740

gtt ctt gtc atg cct acg ttc cat gtc cca ttt aca gat ctt cag gtt     11402
Val Leu Val Met Pro Thr Phe His Val Pro Phe Thr Asp Leu Gln Val
        3745                3750                3755

cca tcg tgc aaa ctt gac ttc aga gaa ata caa atc tat aag aag ctg     11450
Pro Ser Cys Lys Leu Asp Phe Arg Glu Ile Gln Ile Tyr Lys Lys Leu
    3760                3765                3770

aga act tca tca ttt gcc ctc aac cta cca aca ctc ccc gag gta aaa     11498
Arg Thr Ser Ser Phe Ala Leu Asn Leu Pro Thr Leu Pro Glu Val Lys
3775                3780                3785                3790

ttc cct gaa gtt gat gtg tta aca aaa tat tct caa cca gaa gac tcc     11546
Phe Pro Glu Val Asp Val Leu Thr Lys Tyr Ser Gln Pro Glu Asp Ser
```

```
                3795                    3800                    3805

ttg att ccc ttt ttt gag ata acc gtg cct gaa tct cag tta act gtg     11594
Leu Ile Pro Phe Phe Glu Ile Thr Val Pro Glu Ser Gln Leu Thr Val
            3810                    3815                    3820

tcc cag ttc acg ctt cca aaa agt gtt tca gat ggc att gct gct ttg     11642
Ser Gln Phe Thr Leu Pro Lys Ser Val Ser Asp Gly Ile Ala Ala Leu
        3825                    3830                    3835

gat cta aat gca gta gcc aac aag atc gca gac ttt gag ttg ccc acc     11690
Asp Leu Asn Ala Val Ala Asn Lys Ile Ala Asp Phe Glu Leu Pro Thr
    3840                    3845                    3850

atc atc gtg cct gag cag acc att gag att ccc tcc att aag ttc tct     11738
Ile Ile Val Pro Glu Gln Thr Ile Glu Ile Pro Ser Ile Lys Phe Ser
3855                    3860                    3865                    3870

gta cct gct gga att gtc att cct tcc ttt caa gca ctg act gca cgc     11786
Val Pro Ala Gly Ile Val Ile Pro Ser Phe Gln Ala Leu Thr Ala Arg
                3875                    3880                    3885

ttt gag gta gac tct ccc gtg tat aat gcc act tgg agt gcc agt ttg     11834
Phe Glu Val Asp Ser Pro Val Tyr Asn Ala Thr Trp Ser Ala Ser Leu
            3890                    3895                    3900

aaa aac aaa gca gat tat gtt gaa aca gtc ctg gat tcc aca tgc agc     11882
Lys Asn Lys Ala Asp Tyr Val Glu Thr Val Leu Asp Ser Thr Cys Ser
        3905                    3910                    3915

tca acc gta cag ttc cta gaa tat gaa cta aat gtt ttg gga aca cac     11930
Ser Thr Val Gln Phe Leu Glu Tyr Glu Leu Asn Val Leu Gly Thr His
    3920                    3925                    3930

aaa atc gaa gat ggt acg tta gcc tct aag act aaa gga aca ctt gca     11978
Lys Ile Glu Asp Gly Thr Leu Ala Ser Lys Thr Lys Gly Thr Leu Ala
3935                    3940                    3945                    3950

cac cgt gac ttc agt gca gaa tat gaa gaa gat ggc aaa ttt gaa gga     12026
His Arg Asp Phe Ser Ala Glu Tyr Glu Glu Asp Gly Lys Phe Glu Gly
                3955                    3960                    3965

ctt cag gaa tgg gaa gga aaa gcg cac ctc aat atc aaa agc cca gcg     12074
Leu Gln Glu Trp Glu Gly Lys Ala His Leu Asn Ile Lys Ser Pro Ala
            3970                    3975                    3980

ttc acc gat ctc cat ctg cgc tac cag aaa gac aag aaa ggc atc tcc     12122
Phe Thr Asp Leu His Leu Arg Tyr Gln Lys Asp Lys Lys Gly Ile Ser
        3985                    3990                    3995

acc tca gca gcc tcc cca gcc gta ggc acc gtg ggc atg gat atg gat     12170
Thr Ser Ala Ala Ser Pro Ala Val Gly Thr Val Gly Met Asp Met Asp
    4000                    4005                    4010

gaa gat gac gac ttt tct aaa tgg aac ttc tac tac agc cct cag tcc     12218
Glu Asp Asp Asp Phe Ser Lys Trp Asn Phe Tyr Tyr Ser Pro Gln Ser
4015                    4020                    4025                    4030
```

```
tct cca gat aaa aaa ctc acc ata ttc aaa act gag ttg agg gtc cgg    12266
Ser Pro Asp Lys Lys Leu Thr Ile Phe Lys Thr Glu Leu Arg Val Arg
                4035                4040                4045

gaa tct gat gag gaa act cag atc aaa gtt aat tgg gaa gaa gag gca    12314
Glu Ser Asp Glu Glu Thr Gln Ile Lys Val Asn Trp Glu Glu Glu Ala
            4050                4055                4060

gct tct ggc ttg cta acc tct ctg aaa gac aac gtg ccc aag gcc aca    12362
Ala Ser Gly Leu Leu Thr Ser Leu Lys Asp Asn Val Pro Lys Ala Thr
        4065                4070                4075

ggg gtc ctt tat gat tat gtc aac aag tac cac tgg gaa cac aca ggg    12410
Gly Val Leu Tyr Asp Tyr Val Asn Lys Tyr His Trp Glu His Thr Gly
    4080                4085                4090

ctc acc ctg aga gaa gtg tct tca aag ctg aga aga aat ctg cag aac    12458
Leu Thr Leu Arg Glu Val Ser Ser Lys Leu Arg Arg Asn Leu Gln Asn
4095                4100                4105                4110

aat gct gag tgg gtt tat caa ggg gcc att agg caa att gat gat atc    12506
Asn Ala Glu Trp Val Tyr Gln Gly Ala Ile Arg Gln Ile Asp Asp Ile
                4115                4120                4125

gac gtg agg ttc cag aaa gca gcc agt ggc acc act ggg acc tac caa    12554
Asp Val Arg Phe Gln Lys Ala Ala Ser Gly Thr Thr Gly Thr Tyr Gln
            4130                4135                4140

gag tgg aag gac aag gcc cag aat ctg tac cag gaa ctg ttg act cag    12602
Glu Trp Lys Asp Lys Ala Gln Asn Leu Tyr Gln Glu Leu Leu Thr Gln
        4145                4150                4155

gaa ggc caa gcc agt ttc cag gga ctc aag gat aac gtg ttt gat ggc    12650
Glu Gly Gln Ala Ser Phe Gln Gly Leu Lys Asp Asn Val Phe Asp Gly
    4160                4165                4170

ttg gta cga gtt act caa aaa ttc cat atg aaa gtc aag cat ctg att    12698
Leu Val Arg Val Thr Gln Lys Phe His Met Lys Val Lys His Leu Ile
4175                4180                4185                4190

gac tca ctc att gat ttt ctg aac ttc ccc aga ttc cag ttt ccg ggg    12746
Asp Ser Leu Ile Asp Phe Leu Asn Phe Pro Arg Phe Gln Phe Pro Gly
                4195                4200                4205

aaa cct ggg ata tac act agg gag gaa ctt tgc act atg ttc ata agg    12794
Lys Pro Gly Ile Tyr Thr Arg Glu Glu Leu Cys Thr Met Phe Ile Arg
            4210                4215                4220

gag gta ggg acg gta ctg tcc cag gta tat tcg aaa gtc cat aat ggt    12842
Glu Val Gly Thr Val Leu Ser Gln Val Tyr Ser Lys Val His Asn Gly
        4225                4230                4235

tca gaa ata ctg ttt tcc tat ttc caa gac cta gtg att aca ctt cct    12890
Ser Glu Ile Leu Phe Ser Tyr Phe Gln Asp Leu Val Ile Thr Leu Pro
    4240                4245                4250

ttc gag tta agg aaa cat aaa cta ata gat gta atc tcg atg tat agg    12938
Phe Glu Leu Arg Lys His Lys Leu Ile Asp Val Ile Ser Met Tyr Arg
```

```
4255                4260                4265                4270

gaa ctg ttg aaa gat tta tca aaa gaa gcc caa gag gta ttt aaa gcc     12986
Glu Leu Leu Lys Asp Leu Ser Lys Glu Ala Gln Glu Val Phe Lys Ala
                4275                4280                4285

att cag tct ctc aag acc aca gag gtg cta cgt aat ctt cag gac ctt     13034
Ile Gln Ser Leu Lys Thr Thr Glu Val Leu Arg Asn Leu Gln Asp Leu
            4290                4295                4300

tta caa ttc att ttc caa cta ata gaa gat aac att aaa cag ctg aaa     13082
Leu Gln Phe Ile Phe Gln Leu Ile Glu Asp Asn Ile Lys Gln Leu Lys
        4305                4310                4315

gag atg aaa ttt act tat ctt att aat tat atc caa gat gag atc aac     13130
Glu Met Lys Phe Thr Tyr Leu Ile Asn Tyr Ile Gln Asp Glu Ile Asn
    4320                4325                4330

aca atc ttc aat gat tat atc cca tat gtt ttt aaa ttg ttg aaa gaa     13178
Thr Ile Phe Asn Asp Tyr Ile Pro Tyr Val Phe Lys Leu Leu Lys Glu
4335                4340                4345                4350

aac cta tgc ctt aat ctt cat aag ttc aat gaa ttt att caa aac gag     13226
Asn Leu Cys Leu Asn Leu His Lys Phe Asn Glu Phe Ile Gln Asn Glu
                4355                4360                4365

ctt cag gaa gct tct caa gag tta cag cag atc cat caa tac att atg     13274
Leu Gln Glu Ala Ser Gln Glu Leu Gln Gln Ile His Gln Tyr Ile Met
            4370                4375                4380

gcc ctt cgt gaa gaa tat ttt gat cca agt ata gtt ggc tgg aca gtg     13322
Ala Leu Arg Glu Glu Tyr Phe Asp Pro Ser Ile Val Gly Trp Thr Val
        4385                4390                4395

aaa tat tat gaa ctt gaa gaa aag ata gtc agt ctg atc aag aac ctg     13370
Lys Tyr Tyr Glu Leu Glu Glu Lys Ile Val Ser Leu Ile Lys Asn Leu
    4400                4405                4410

tta gtt gct ctt aag gac ttc cat tct gaa tat att gtc agt gcc tct     13418
Leu Val Ala Leu Lys Asp Phe His Ser Glu Tyr Ile Val Ser Ala Ser
4415                4420                4425                4430

aac ttt act tcc caa ctc tca agt caa gtt gag caa ttt ctg cac aga     13466
Asn Phe Thr Ser Gln Leu Ser Ser Gln Val Glu Gln Phe Leu His Arg
                4435                4440                4445

aat att cag gaa tat ctt agc atc ctt acc gat cca gat gga aaa ggg     13514
Asn Ile Gln Glu Tyr Leu Ser Ile Leu Thr Asp Pro Asp Gly Lys Gly
            4450                4455                4460

aaa gag aag att gca gag ctt tct gcc act gct cag gaa ata att aaa     13562
Lys Glu Lys Ile Ala Glu Leu Ser Ala Thr Ala Gln Glu Ile Ile Lys
        4465                4470                4475

agc cag gcc att gcg acg aag aaa ata att tct gat tac cac cag cag     13610
Ser Gln Ala Ile Ala Thr Lys Lys Ile Ile Ser Asp Tyr His Gln Gln
    4480                4485                4490
```

```
ttt aga tat aaa ctg caa gat ttt tca gac caa ctc tct gat tac tat    13658
Phe Arg Tyr Lys Leu Gln Asp Phe Ser Asp Gln Leu Ser Asp Tyr Tyr
4495                4500                4505                4510

gaa aaa ttt att gct gaa tcc aaa aga ttg att gac ctg tcc att caa    13706
Glu Lys Phe Ile Ala Glu Ser Lys Arg Leu Ile Asp Leu Ser Ile Gln
                4515                4520                4525

aac tac cac aca ttt ctg ata tac atc acg gag tta ctg aaa aag ctg    13754
Asn Tyr His Thr Phe Leu Ile Tyr Ile Thr Glu Leu Leu Lys Lys Leu
            4530                4535                4540

caa tca acc aca gtc atg aac ccc tac atg aag ctt gct cca gga gaa    13802
Gln Ser Thr Thr Val Met Asn Pro Tyr Met Lys Leu Ala Pro Gly Glu
        4545                4550                4555

ctt act atc atc ctc taa tttttaaaa gaaatcttca tttattcttc           13850
Leu Thr Ile Ile Leu  *
    4560

ttttccaatt gaactttcac atagcacaga aaaaattcaa actgcctata ttgataaaac  13910
catacagtga gccagccttg cagtaggcag tagactataa gcagaagcac atatgaactg  13970
gacctgcacc aaagctggca ccagggctcg gaaggtctct gaactcagaa ggatggcatt  14030
ttttgcaagt taaagaaaat caggatctga gttattttgc taaacttggg ggaggaggaa  14090
caaataaatg gagtctttat tgtgtatcat a                                 14121

&lt;210&gt; 32
&lt;211&gt; 4563
&lt;212&gt; PRT
&lt;213&gt; Homo sapien

&lt;400&gt; 32
Met Asp Pro Pro Arg Pro Ala Leu Leu Ala Leu Leu Ala Leu Pro Ala
 1               5                  10                  15
Leu Leu Leu Leu Leu Leu Ala Gly Ala Arg Ala Glu Glu Glu Met Leu
            20                  25                  30
Glu Asn Val Ser Leu Val Cys Pro Lys Asp Ala Thr Arg Phe Lys His
        35                  40                  45
Leu Arg Lys Tyr Thr Tyr Asn Tyr Glu Ala Glu Ser Ser Ser Gly Val
    50                  55                  60
Pro Gly Thr Ala Asp Ser Arg Ser Ala Thr Arg Ile Asn Cys Lys Val
65                  70                  75                  80
Glu Leu Glu Val Pro Gln Leu Cys Ser Phe Ile Leu Lys Thr Ser Gln
                85                  90                  95
Cys Thr Leu Lys Glu Val Tyr Gly Phe Asn Pro Glu Gly Lys Ala Leu
            100                 105                 110
Leu Lys Lys Thr Lys Asn Ser Glu Glu Phe Ala Ala Ala Met Ser Arg
        115                 120                 125
Tyr Glu Leu Lys Leu Ala Ile Pro Glu Gly Lys Gln Val Phe Leu Tyr
    130                 135                 140
Pro Glu Lys Asp Glu Pro Thr Tyr Ile Leu Asn Ile Lys Arg Gly Ile
145                 150                 155                 160
Ile Ser Ala Leu Leu Val Pro Pro Glu Thr Glu Glu Ala Lys Gln Val
                165                 170                 175
Leu Phe Leu Asp Thr Val Tyr Gly Asn Cys Ser Thr His Phe Thr Val
            180                 185                 190
Lys Thr Arg Lys Gly Asn Val Ala Thr Glu Ile Ser Thr Glu Arg Asp
        195                 200                 205
```

```
Leu Gly Gln Cys Asp Arg Phe Lys Pro Ile Arg Thr Gly Ile Ser Pro
    210                 215                 220
Leu Ala Leu Ile Lys Gly Met Thr Arg Pro Leu Ser Thr Leu Ile Ser
225                 230                 235                 240
Ser Ser Gln Ser Cys Gln Tyr Thr Leu Asp Ala Lys Arg Lys His Val
                245                 250                 255
Ala Glu Ala Ile Cys Lys Glu Gln His Leu Phe Leu Pro Phe Ser Tyr
            260                 265                 270
Asn Asn Lys Tyr Gly Met Val Ala Gln Val Thr Gln Thr Leu Lys Leu
        275                 280                 285
Glu Asp Thr Pro Lys Ile Asn Ser Arg Phe Phe Gly Glu Gly Thr Lys
    290                 295                 300
Lys Met Gly Leu Ala Phe Glu Ser Thr Lys Ser Thr Ser Pro Pro Lys
305                 310                 315                 320
Gln Ala Glu Ala Val Leu Lys Thr Leu Gln Glu Leu Lys Lys Leu Thr
                325                 330                 335
Ile Ser Glu Gln Asn Ile Gln Arg Ala Asn Leu Phe Asn Lys Leu Val
            340                 345                 350
Thr Glu Leu Arg Gly Leu Ser Asp Glu Ala Val Thr Ser Leu Leu Pro
        355                 360                 365
Gln Leu Ile Glu Val Ser Ser Pro Ile Thr Leu Gln Ala Leu Val Gln
    370                 375                 380
Cys Gly Gln Pro Gln Cys Ser Thr His Ile Leu Gln Trp Leu Lys Arg
385                 390                 395                 400
Val His Ala Asn Pro Leu Leu Ile Asp Val Val Thr Tyr Leu Val Ala
                405                 410                 415
Leu Ile Pro Glu Pro Ser Ala Gln Gln Leu Arg Glu Ile Phe Asn Met
            420                 425                 430
Ala Arg Asp Gln Arg Ser Arg Ala Thr Leu Tyr Ala Leu Ser His Ala
        435                 440                 445
Val Asn Asn Tyr His Lys Thr Asn Pro Thr Gly Thr Gln Glu Leu Leu
    450                 455                 460
Asp Ile Ala Asn Tyr Leu Met Glu Gln Ile Gln Asp Asp Cys Thr Gly
465                 470                 475                 480
Asp Glu Asp Tyr Thr Tyr Leu Ile Leu Arg Val Ile Gly Asn Met Gly
                485                 490                 495
Gln Thr Met Glu Gln Leu Thr Pro Glu Leu Lys Ser Ser Ile Leu Lys
            500                 505                 510
Cys Val Gln Ser Thr Lys Pro Ser Leu Met Ile Gln Lys Ala Ala Ile
        515                 520                 525
Gln Ala Leu Arg Lys Met Glu Pro Lys Asp Lys Asp Gln Glu Val Leu
    530                 535                 540
Leu Gln Thr Phe Leu Asp Asp Ala Ser Pro Gly Asp Lys Arg Leu Ala
545                 550                 555                 560
Ala Tyr Leu Met Leu Met Arg Ser Pro Ser Gln Ala Asp Ile Asn Lys
                565                 570                 575
Ile Val Gln Ile Leu Pro Trp Glu Gln Asn Glu Gln Val Lys Asn Phe
            580                 585                 590
Val Ala Ser His Ile Ala Asn Ile Leu Asn Ser Glu Glu Leu Asp Ile
        595                 600                 605
Gln Asp Leu Lys Lys Leu Val Lys Glu Ala Leu Lys Glu Ser Gln Leu
    610                 615                 620
Pro Thr Val Met Asp Phe Arg Lys Phe Ser Arg Asn Tyr Gln Leu Tyr
625                 630                 635                 640
Lys Ser Val Ser Leu Pro Ser Leu Asp Pro Ala Ser Ala Lys Ile Glu
                645                 650                 655
Gly Asn Leu Ile Phe Asp Pro Asn Asn Tyr Leu Pro Lys Glu Ser Met
            660                 665                 670
```

```
Leu Lys Thr Thr Leu Thr Ala Phe Gly Phe Ala Ser Ala Asp Leu Ile
        675                 680                 685
Glu Ile Gly Leu Glu Gly Lys Gly Phe Glu Pro Thr Leu Glu Ala Leu
    690                 695                 700
Phe Gly Lys Gln Gly Phe Phe Pro Asp Ser Val Asn Lys Ala Leu Tyr
705                 710                 715                 720
Trp Val Asn Gly Gln Val Pro Asp Gly Val Ser Lys Val Leu Val Asp
                725                 730                 735
His Phe Gly Tyr Thr Lys Asp Asp Lys His Glu Gln Asp Met Val Asn
            740                 745                 750
Gly Ile Met Leu Ser Val Glu Lys Leu Ile Lys Asp Leu Lys Ser Lys
        755                 760                 765
Glu Val Pro Glu Ala Arg Ala Tyr Leu Arg Ile Leu Gly Glu Glu Leu
    770                 775                 780
Gly Phe Ala Ser Leu His Asp Leu Gln Leu Leu Gly Lys Leu Leu Leu
785                 790                 795                 800
Met Gly Ala Arg Thr Leu Gln Gly Ile Pro Gln Met Ile Gly Glu Val
                805                 810                 815
Ile Arg Lys Gly Ser Lys Asn Asp Phe Phe Leu His Tyr Ile Phe Met
            820                 825                 830
Glu Asn Ala Phe Glu Leu Pro Thr Gly Ala Gly Leu Gln Leu Gln Ile
        835                 840                 845
Ser Ser Ser Gly Val Ile Ala Pro Gly Ala Lys Ala Gly Val Lys Leu
    850                 855                 860
Glu Val Ala Asn Met Gln Ala Glu Leu Val Ala Lys Pro Ser Val Ser
865                 870                 875                 880
Val Glu Phe Val Thr Asn Met Gly Ile Ile Pro Asp Phe Ala Arg
                885                 890                 895
Ser Gly Val Gln Met Asn Thr Asn Phe Phe His Glu Ser Gly Leu Glu
            900                 905                 910
Ala His Val Ala Leu Lys Ala Gly Lys Leu Lys Phe Ile Ile Pro Ser
        915                 920                 925
Pro Lys Arg Pro Val Lys Leu Leu Ser Gly Gly Asn Thr Leu His Leu
    930                 935                 940
Val Ser Thr Thr Lys Thr Glu Val Ile Pro Pro Leu Ile Glu Asn Arg
945                 950                 955                 960
Gln Ser Trp Ser Val Cys Lys Gln Val Phe Pro Gly Leu Asn Tyr Cys
                965                 970                 975
Thr Ser Gly Ala Tyr Ser Asn Ala Ser Ser Thr Asp Ser Ala Ser Tyr
            980                 985                 990
Tyr Pro Leu Thr Gly Asp Thr Arg Leu Glu Leu Glu Leu Arg Pro Thr
        995                 1000                1005
Gly Glu Ile Glu Gln Tyr Ser Val Ser Ala Thr Tyr Glu Leu Gln Arg
    1010                1015                1020
Glu Asp Arg Ala Leu Val Asp Thr Leu Lys Phe Val Thr Gln Ala Glu
1025                1030                1035                1040
Gly Ala Lys Gln Thr Glu Ala Thr Met Thr Phe Lys Tyr Asn Arg Gln
                1045                1050                1055
Ser Met Thr Leu Ser Ser Glu Val Gln Ile Pro Asp Phe Asp Val Asp
            1060                1065                1070
Leu Gly Thr Ile Leu Arg Val Asn Asp Glu Ser Thr Glu Gly Lys Thr
        1075                1080                1085
Ser Tyr Arg Leu Thr Leu Asp Ile Gln Asn Lys Lys Ile Thr Glu Val
    1090                1095                1100
Ala Leu Met Gly His Leu Ser Cys Asp Thr Lys Glu Glu Arg Lys Ile
1105                1110                1115                1120
Lys Gly Val Ile Ser Ile Pro Arg Leu Gln Ala Glu Ala Arg Ser Glu
                1125                1130                1135
```

```
Ile Leu Ala His Trp Ser Pro Ala Lys Leu Leu Leu Gln Met Asp Ser
            1140                1145                1150
Ser Ala Thr Ala Tyr Gly Ser Thr Val Ser Lys Arg Val Ala Trp His
        1155                1160                1165
Tyr Asp Glu Glu Lys Ile Glu Phe Glu Trp Asn Thr Gly Thr Asn Val
    1170                1175                1180
Asp Thr Lys Lys Met Thr Ser Asn Phe Pro Val Asp Leu Ser Asp Tyr
1185                1190                1195                1200
Pro Lys Ser Leu His Met Tyr Ala Asn Arg Leu Leu Asp His Arg Val
                1205                1210                1215
Pro Glu Thr Asp Met Thr Phe Arg His Val Gly Ser Lys Leu Ile Val
            1220                1225                1230
Ala Met Ser Ser Trp Leu Gln Lys Ala Ser Gly Ser Leu Pro Tyr Thr
        1235                1240                1245
Gln Thr Leu Gln Asp His Leu Asn Ser Leu Lys Glu Phe Asn Leu Gln
    1250                1255                1260
Asn Met Gly Leu Pro Asp Phe His Ile Pro Glu Asn Leu Phe Leu Lys
1265                1270                1275                1280
Ser Asp Gly Arg Val Lys Tyr Thr Leu Asn Lys Asn Ser Leu Lys Ile
                1285                1290                1295
Glu Ile Pro Leu Pro Phe Gly Gly Lys Ser Ser Arg Asp Leu Lys Met
            1300                1305                1310
Leu Glu Thr Val Arg Thr Pro Ala Leu His Phe Lys Ser Val Gly Phe
        1315                1320                1325
His Leu Pro Ser Arg Glu Phe Gln Val Pro Thr Phe Thr Ile Pro Lys
    1330                1335                1340
Leu Tyr Gln Leu Gln Val Pro Leu Leu Gly Val Leu Asp Leu Ser Thr
1345                1350                1355                1360
Asn Val Tyr Ser Asn Leu Tyr Asn Trp Ser Ala Ser Tyr Ser Gly Gly
                1365                1370                1375
Asn Thr Ser Thr Asp His Phe Ser Leu Arg Ala Arg Tyr His Met Lys
            1380                1385                1390
Ala Asp Ser Val Val Asp Leu Leu Ser Tyr Asn Val Gln Gly Ser Gly
        1395                1400                1405
Glu Thr Thr Tyr Asp His Lys Asn Thr Phe Thr Leu Ser Cys Asp Gly
    1410                1415                1420
Ser Leu Arg His Lys Phe Leu Asp Ser Asn Ile Lys Phe Ser His Val
1425                1430                1435                1440
Glu Lys Leu Gly Asn Asn Pro Val Ser Lys Gly Leu Leu Ile Phe Asp
                1445                1450                1455
Ala Ser Ser Ser Trp Gly Pro Gln Met Ser Ala Ser Val His Leu Asp
            1460                1465                1470
Ser Lys Lys Lys Gln His Leu Phe Val Lys Glu Val Lys Ile Asp Gly
        1475                1480                1485
Gln Phe Arg Val Ser Ser Phe Tyr Ala Lys Gly Thr Tyr Gly Leu Ser
    1490                1495                1500
Cys Gln Arg Asp Pro Asn Thr Gly Arg Leu Asn Gly Glu Ser Asn Leu
1505                1510                1515                1520
Arg Phe Asn Ser Ser Tyr Leu Gln Gly Thr Asn Gln Ile Thr Gly Arg
                1525                1530                1535
Tyr Glu Asp Gly Thr Leu Ser Leu Thr Ser Thr Ser Asp Leu Gln Ser
            1540                1545                1550
Gly Ile Ile Lys Asn Thr Ala Ser Leu Lys Tyr Glu Asn Tyr Glu Leu
        1555                1560                1565
Thr Leu Lys Ser Asp Thr Asn Gly Lys Tyr Lys Asn Phe Ala Thr Ser
    1570                1575                1580
Asn Lys Met Asp Met Thr Phe Ser Lys Gln Asn Ala Leu Leu Arg Ser
1585                1590                1595                1600
```

```
Glu Tyr Gln Ala Asp Tyr Glu Ser Leu Arg Phe Phe Ser Leu Leu Ser
                1605                1610                1615
Gly Ser Leu Asn Ser His Gly Leu Glu Leu Asn Ala Asp Ile Leu Gly
            1620                1625                1630
Thr Asp Lys Ile Asn Ser Gly Ala His Lys Ala Thr Leu Arg Ile Gly
        1635                1640                1645
Gln Asp Gly Ile Ser Thr Ser Ala Thr Thr Asn Leu Lys Cys Ser Leu
    1650                1655                1660
Leu Val Leu Glu Asn Glu Leu Asn Ala Glu Leu Gly Leu Ser Gly Ala
1665                1670                1675                1680
Ser Met Lys Leu Thr Thr Asn Gly Arg Phe Arg Glu His Asn Ala Lys
                1685                1690                1695
Phe Ser Leu Asp Gly Lys Ala Ala Leu Thr Glu Leu Ser Leu Gly Ser
            1700                1705                1710
Ala Tyr Gln Ala Met Ile Leu Gly Val Asp Ser Lys Asn Ile Phe Asn
        1715                1720                1725
Phe Lys Val Ser Gln Glu Gly Leu Lys Leu Ser Asn Asp Met Met Gly
    1730                1735                1740
Ser Tyr Ala Glu Met Lys Phe Asp His Thr Asn Ser Leu Asn Ile Ala
1745                1750                1755                1760
Gly Leu Ser Leu Asp Phe Ser Ser Lys Leu Asp Asn Ile Tyr Ser Ser
                1765                1770                1775
Asp Lys Phe Tyr Lys Gln Thr Val Asn Leu Gln Leu Gln Pro Tyr Ser
            1780                1785                1790
Leu Val Thr Thr Leu Asn Ser Asp Leu Lys Tyr Asn Ala Leu Asp Leu
        1795                1800                1805
Thr Asn Asn Gly Lys Leu Arg Leu Glu Pro Leu Lys Leu His Val Ala
    1810                1815                1820
Gly Asn Leu Lys Gly Ala Tyr Gln Asn Asn Glu Ile Lys His Ile Tyr
1825                1830                1835                1840
Ala Ile Ser Ser Ala Ala Leu Ser Ala Ser Tyr Lys Ala Asp Thr Val
                1845                1850                1855
Ala Lys Val Gln Gly Val Glu Phe Ser His Arg Leu Asn Thr Asp Ile
            1860                1865                1870
Ala Gly Leu Ala Ser Ala Ile Asp Met Ser Thr Asn Tyr Asn Ser Asp
        1875                1880                1885
Ser Leu His Phe Ser Asn Val Phe Arg Ser Val Met Ala Pro Phe Thr
    1890                1895                1900
Met Thr Ile Asp Ala His Thr Asn Gly Asn Gly Lys Leu Ala Leu Trp
1905                1910                1915                1920
Gly Glu His Thr Gly Gln Leu Tyr Ser Lys Phe Leu Leu Lys Ala Glu
                1925                1930                1935
Pro Leu Ala Phe Thr Phe Ser His Asp Tyr Lys Gly Ser Thr Ser His
            1940                1945                1950
His Leu Val Ser Arg Lys Ser Ile Ser Ala Ala Leu Glu His Lys Val
        1955                1960                1965
Ser Ala Leu Leu Thr Pro Ala Glu Gln Thr Gly Thr Trp Lys Leu Lys
    1970                1975                1980
Thr Gln Phe Asn Asn Asn Glu Tyr Ser Gln Asp Leu Asp Ala Tyr Asn
1985                1990                1995                2000
Thr Lys Asp Lys Ile Gly Val Glu Leu Thr Gly Arg Thr Leu Ala Asp
                2005                2010                2015
Leu Thr Leu Leu Asp Ser Pro Ile Lys Val Pro Leu Leu Leu Ser Glu
            2020                2025                2030
Pro Ile Asn Ile Ile Asp Ala Leu Glu Met Arg Asp Ala Val Glu Lys
        2035                2040                2045
Pro Gln Glu Phe Thr Ile Val Ala Phe Val Lys Tyr Asp Lys Asn Gln
    2050                2055                2060
```

```
Asp Val His Ser Ile Asn Leu Pro Phe Phe Glu Thr Leu Gln Glu Tyr
2065                2070                2075                2080
Phe Glu Arg Asn Arg Gln Thr Ile Ile Val Val Val Glu Asn Val Gln
                2085                2090                2095
Arg Asn Leu Lys His Ile Asn Ile Asp Gln Phe Val Arg Lys Tyr Arg
            2100                2105                2110
Ala Ala Leu Gly Lys Leu Pro Gln Gln Ala Asn Asp Tyr Leu Asn Ser
        2115                2120                2125
Phe Asn Trp Glu Arg Gln Val Ser His Ala Lys Glu Lys Leu Thr Ala
    2130                2135                2140
Leu Thr Lys Lys Tyr Arg Ile Thr Glu Asn Asp Ile Gln Ile Ala Leu
2145                2150                2155                2160
Asp Asp Ala Lys Ile Asn Phe Asn Glu Lys Leu Ser Gln Leu Gln Thr
                2165                2170                2175
Tyr Met Ile Gln Phe Asp Gln Tyr Ile Lys Asp Ser Tyr Asp Leu His
            2180                2185                2190
Asp Leu Lys Ile Ala Ile Ala Asn Ile Ile Asp Glu Ile Ile Glu Lys
        2195                2200                2205
Leu Lys Ser Leu Asp Glu His Tyr His Ile Arg Val Asn Leu Val Lys
    2210                2215                2220
Thr Ile His Asp Leu His Leu Phe Ile Glu Asn Ile Asp Phe Asn Lys
2225                2230                2235                2240
Ser Gly Ser Ser Thr Ala Ser Trp Ile Gln Asn Val Asp Thr Lys Tyr
                2245                2250                2255
Gln Ile Arg Ile Gln Ile Gln Glu Lys Leu Gln Gln Leu Lys Arg His
            2260                2265                2270
Ile Gln Asn Ile Asp Ile Gln His Leu Ala Gly Lys Leu Lys Gln His
        2275                2280                2285
Ile Glu Ala Ile Asp Val Arg Val Leu Leu Asp Gln Leu Gly Thr Thr
    2290                2295                2300
Ile Ser Phe Glu Arg Ile Asn Asp Val Leu Glu His Val Lys His Phe
2305                2310                2315                2320
Val Ile Asn Leu Ile Gly Asp Phe Glu Val Ala Glu Lys Ile Asn Ala
                2325                2330                2335
Phe Arg Ala Lys Val His Glu Leu Ile Glu Arg Tyr Glu Val Asp Gln
            2340                2345                2350
Gln Ile Gln Val Leu Met Asp Lys Leu Val Glu Leu Thr His Gln Tyr
        2355                2360                2365
Lys Leu Lys Glu Thr Ile Gln Lys Leu Ser Asn Val Leu Gln Gln Val
    2370                2375                2380
Lys Ile Lys Asp Tyr Phe Glu Lys Leu Val Gly Phe Ile Asp Asp Ala
2385                2390                2395                2400
Val Lys Lys Leu Asn Glu Leu Ser Phe Lys Thr Phe Ile Glu Asp Val
                2405                2410                2415
Asn Lys Phe Leu Asp Met Leu Ile Lys Lys Leu Lys Ser Phe Asp Tyr
            2420                2425                2430
His Gln Phe Val Asp Glu Thr Asn Asp Lys Ile Arg Glu Val Thr Gln
        2435                2440                2445
Arg Leu Asn Gly Glu Ile Gln Ala Leu Glu Leu Pro Gln Lys Ala Glu
    2450                2455                2460
Ala Leu Lys Leu Phe Leu Glu Glu Thr Lys Ala Thr Val Ala Val Tyr
2465                2470                2475                2480
Leu Glu Ser Leu Gln Asp Thr Lys Ile Thr Leu Ile Ile Asn Trp Leu
                2485                2490                2495
Gln Glu Ala Leu Ser Ser Ala Ser Leu Ala His Met Lys Ala Lys Phe
            2500                2505                2510
Arg Glu Thr Leu Glu Asp Thr Arg Asp Arg Met Tyr Gln Met Asp Ile
        2515                2520                2525
```

```
Gln Gln Glu Leu Gln Arg Tyr Leu Ser Leu Val Gly Gln Val Tyr Ser
    2530                2535                2540
Thr Leu Val Thr Tyr Ile Ser Asp Trp Trp Thr Leu Ala Ala Lys Asn
2545                2550                2555                2560
Leu Thr Asp Phe Ala Glu Gln Tyr Ser Ile Gln Asp Trp Ala Lys Arg
                2565                2570                2575
Met Lys Ala Leu Val Glu Gln Gly Phe Thr Val Pro Glu Ile Lys Thr
            2580                2585                2590
Ile Leu Gly Thr Met Pro Ala Phe Glu Val Ser Leu Gln Ala Leu Gln
        2595                2600                2605
Lys Ala Thr Phe Gln Thr Pro Asp Phe Ile Val Pro Leu Thr Asp Leu
    2610                2615                2620
Arg Ile Pro Ser Val Gln Ile Asn Phe Lys Asp Leu Lys Asn Ile Lys
2625                2630                2635                2640
Ile Pro Ser Arg Phe Ser Thr Pro Glu Phe Thr Ile Leu Asn Thr Phe
                2645                2650                2655
His Ile Pro Ser Phe Thr Ile Asp Phe Val Glu Met Lys Val Lys Ile
            2660                2665                2670
Ile Arg Thr Ile Asp Gln Met Gln Asn Ser Glu Leu Gln Trp Pro Val
        2675                2680                2685
Pro Asp Ile Tyr Leu Arg Asp Leu Lys Val Glu Asp Ile Pro Leu Ala
    2690                2695                2700
Arg Ile Thr Leu Pro Asp Phe Arg Leu Pro Glu Ile Ala Ile Pro Glu
2705                2710                2715                2720
Phe Ile Ile Pro Thr Leu Asn Leu Asn Asp Phe Gln Val Pro Asp Leu
                2725                2730                2735
His Ile Pro Glu Phe Gln Leu Pro His Ile Ser His Thr Ile Glu Val
            2740                2745                2750
Pro Thr Phe Gly Lys Leu Tyr Ser Ile Leu Lys Ile Gln Ser Pro Leu
        2755                2760                2765
Phe Thr Leu Asp Ala Asn Ala Asp Ile Gly Asn Gly Thr Thr Ser Ala
    2770                2775                2780
Asn Glu Ala Gly Ile Ala Ala Ser Ile Thr Ala Lys Gly Glu Ser Lys
2785                2790                2795                2800
Leu Glu Val Leu Asn Phe Asp Phe Gln Ala Asn Ala Gln Leu Ser Asn
                2805                2810                2815
Pro Lys Ile Asn Pro Leu Ala Leu Lys Glu Ser Val Lys Phe Ser Ser
            2820                2825                2830
Lys Tyr Leu Arg Thr Glu His Gly Ser Glu Met Leu Phe Phe Gly Asn
        2835                2840                2845
Ala Ile Glu Gly Lys Ser Asn Thr Val Ala Ser Leu His Thr Glu Lys
    2850                2855                2860
Asn Thr Leu Glu Leu Ser Asn Gly Val Ile Val Lys Ile Asn Asn Gln
2865                2870                2875                2880
Leu Thr Leu Asp Ser Asn Thr Lys Tyr Phe His Lys Leu Asn Ile Pro
                2885                2890                2895
Lys Leu Asp Phe Ser Ser Gln Ala Asp Leu Arg Asn Glu Ile Lys Thr
            2900                2905                2910
Leu Leu Lys Ala Gly His Ile Ala Trp Thr Ser Ser Gly Lys Gly Ser
        2915                2920                2925
Trp Lys Trp Ala Cys Pro Arg Phe Ser Asp Glu Gly Thr His Glu Ser
    2930                2935                2940
Gln Ile Ser Phe Thr Ile Glu Gly Pro Leu Thr Ser Phe Gly Leu Ser
2945                2950                2955                2960
Asn Lys Ile Asn Ser Lys His Leu Arg Val Asn Gln Asn Leu Val Tyr
                2965                2970                2975
Glu Ser Gly Ser Leu Asn Phe Ser Lys Leu Glu Ile Gln Ser Gln Val
            2980                2985                2990
```

```
Asp Ser Gln His Val Gly His Ser Val Leu Thr Ala Lys Gly Met Ala
        2995                3000                3005
Leu Phe Gly Glu Gly Lys Ala Glu Phe Thr Gly Arg His Asp Ala His
    3010                3015                3020
Leu Asn Gly Lys Val Ile Gly Thr Leu Lys Asn Ser Leu Phe Phe Ser
3025                3030                3035                3040
Ala Gln Pro Phe Glu Ile Thr Ala Ser Thr Asn Asn Glu Gly Asn Leu
                3045                3050                3055
Lys Val Arg Phe Pro Leu Arg Leu Thr Gly Lys Ile Asp Phe Leu Asn
            3060                3065                3070
Asn Tyr Ala Leu Phe Leu Ser Pro Ser Ala Gln Gln Ala Ser Trp Gln
        3075                3080                3085
Val Ser Ala Arg Phe Asn Gln Tyr Lys Tyr Asn Gln Asn Phe Ser Ala
    3090                3095                3100
Gly Asn Asn Glu Asn Ile Met Glu Ala His Val Gly Ile Asn Gly Glu
3105                3110                3115                3120
Ala Asn Leu Asp Phe Leu Asn Ile Pro Leu Thr Ile Pro Glu Met Arg
                3125                3130                3135
Leu Pro Tyr Thr Ile Ile Thr Thr Pro Pro Leu Lys Asp Phe Ser Leu
            3140                3145                3150
Trp Glu Lys Thr Gly Leu Lys Glu Phe Leu Lys Thr Thr Lys Gln Ser
        3155                3160                3165
Phe Asp Leu Ser Val Lys Ala Gln Tyr Lys Lys Asn Lys His Arg His
    3170                3175                3180
Ser Ile Thr Asn Pro Leu Ala Val Leu Cys Glu Phe Ile Ser Gln Ser
3185                3190                3195                3200
Ile Lys Ser Phe Asp Arg His Phe Glu Lys Asn Arg Asn Asn Ala Leu
                3205                3210                3215
Asp Phe Val Thr Lys Ser Tyr Asn Glu Thr Lys Ile Lys Phe Asp Lys
            3220                3225                3230
Tyr Lys Ala Glu Lys Ser His Asp Glu Leu Pro Arg Thr Phe Gln Ile
        3235                3240                3245
Pro Gly Tyr Thr Val Pro Val Val Asn Val Glu Val Ser Pro Phe Thr
    3250                3255                3260
Ile Glu Met Ser Ala Phe Gly Tyr Val Phe Pro Lys Ala Val Ser Met
3265                3270                3275                3280
Pro Ser Phe Ser Ile Leu Gly Ser Asp Val Arg Val Pro Ser Tyr Thr
                3285                3290                3295
Leu Ile Leu Pro Ser Leu Glu Leu Pro Val Leu His Val Pro Arg Asn
            3300                3305                3310
Leu Lys Leu Ser Leu Pro His Phe Lys Glu Leu Cys Thr Ile Ser His
        3315                3320                3325
Ile Phe Ile Pro Ala Met Gly Asn Ile Thr Tyr Asp Phe Ser Phe Lys
    3330                3335                3340
Ser Ser Val Ile Thr Leu Asn Thr Asn Ala Glu Leu Phe Asn Gln Ser
3345                3350                3355                3360
Asp Ile Val Ala His Leu Leu Ser Ser Ser Ser Ser Val Ile Asp Ala
                3365                3370                3375
Leu Gln Tyr Lys Leu Glu Gly Thr Thr Arg Leu Thr Arg Lys Arg Gly
            3380                3385                3390
Leu Lys Leu Ala Thr Ala Leu Ser Leu Ser Asn Lys Phe Val Glu Gly
        3395                3400                3405
Ser His Asn Ser Thr Val Ser Leu Thr Thr Lys Asn Met Glu Val Ser
    3410                3415                3420
Val Ala Lys Thr Thr Lys Ala Glu Ile Pro Ile Leu Arg Met Asn Phe
3425                3430                3435                3440
Lys Gln Glu Leu Asn Gly Asn Thr Lys Ser Lys Pro Thr Val Ser Ser
                3445                3450                3455
```

```
Ser Met Glu Phe Lys Tyr Asp Phe Asn Ser Ser Met Leu Tyr Ser Thr
            3460                3465                3470
Ala Lys Gly Ala Val Asp His Lys Leu Ser Leu Glu Ser Leu Thr Ser
        3475                3480                3485
Tyr Phe Ser Ile Glu Ser Ser Thr Lys Gly Asp Val Lys Gly Ser Val
    3490                3495                3500
Leu Ser Arg Glu Tyr Ser Gly Thr Ile Ala Ser Glu Ala Asn Thr Tyr
3505                3510                3515                3520
Leu Asn Ser Lys Ser Thr Arg Ser Ser Val Lys Leu Gln Gly Thr Ser
                3525                3530                3535
Lys Ile Asp Asp Ile Trp Asn Leu Glu Val Lys Glu Asn Phe Ala Gly
            3540                3545                3550
Glu Ala Thr Leu Gln Arg Ile Tyr Ser Leu Trp Glu His Ser Thr Lys
        3555                3560                3565
Asn His Leu Gln Leu Glu Gly Leu Phe Phe Thr Asn Gly Glu His Thr
    3570                3575                3580
Ser Lys Ala Thr Leu Glu Leu Ser Pro Trp Gln Met Ser Ala Leu Val
3585                3590                3595                3600
Gln Val His Ala Ser Gln Pro Ser Ser Phe His Asp Phe Pro Asp Leu
                3605                3610                3615
Gly Gln Glu Val Ala Leu Asn Ala Asn Thr Lys Asn Gln Lys Ile Arg
            3620                3625                3630
Trp Lys Asn Glu Val Arg Ile His Ser Gly Ser Phe Gln Ser Gln Val
        3635                3640                3645
Glu Leu Ser Asn Asp Gln Glu Lys Ala His Leu Asp Ile Ala Gly Ser
    3650                3655                3660
Leu Glu Gly His Leu Arg Phe Leu Lys Asn Ile Ile Leu Pro Val Tyr
3665                3670                3675                3680
Asp Lys Ser Leu Trp Asp Phe Leu Lys Leu Asp Val Thr Thr Ser Ile
                3685                3690                3695
Gly Arg Arg Gln His Leu Arg Val Ser Thr Ala Phe Val Tyr Thr Lys
            3700                3705                3710
Asn Pro Asn Gly Tyr Ser Phe Ser Ile Pro Val Lys Val Leu Ala Asp
        3715                3720                3725
Lys Phe Ile Thr Pro Gly Leu Lys Leu Asn Asp Leu Asn Ser Val Leu
    3730                3735                3740
Val Met Pro Thr Phe His Val Pro Phe Thr Asp Leu Gln Val Pro Ser
3745                3750                3755                3760
Cys Lys Leu Asp Phe Arg Glu Ile Gln Ile Tyr Lys Lys Leu Arg Thr
                3765                3770                3775
Ser Ser Phe Ala Leu Asn Leu Pro Thr Leu Pro Glu Val Lys Phe Pro
            3780                3785                3790
Glu Val Asp Val Leu Thr Lys Tyr Ser Gln Pro Glu Asp Ser Leu Ile
        3795                3800                3805
Pro Phe Phe Glu Ile Thr Val Pro Glu Ser Gln Leu Thr Val Ser Gln
    3810                3815                3820
Phe Thr Leu Pro Lys Ser Val Ser Asp Gly Ile Ala Ala Leu Asp Leu
3825                3830                3835                3840
Asn Ala Val Ala Asn Lys Ile Ala Asp Phe Glu Leu Pro Thr Ile Ile
                3845                3850                3855
Val Pro Glu Gln Thr Ile Glu Ile Pro Ser Ile Lys Phe Ser Val Pro
            3860                3865                3870
Ala Gly Ile Val Ile Pro Ser Phe Gln Ala Leu Thr Ala Arg Phe Glu
        3875                3880                3885
Val Asp Ser Pro Val Tyr Asn Ala Thr Trp Ser Ala Ser Leu Lys Asn
    3890                3895                3900
Lys Ala Asp Tyr Val Glu Thr Val Leu Asp Ser Thr Cys Ser Ser Thr
3905                3910                3915                3920
```

```
Val Gln Phe Leu Glu Tyr Glu Leu Asn Val Leu Gly Thr His Lys Ile
                3925                3930                3935
Glu Asp Gly Thr Leu Ala Ser Lys Thr Lys Gly Thr Leu Ala His Arg
            3940                3945                3950
Asp Phe Ser Ala Glu Tyr Glu Glu Asp Gly Lys Phe Glu Gly Leu Gln
        3955                3960                3965
Glu Trp Glu Gly Lys Ala His Leu Asn Ile Lys Ser Pro Ala Phe Thr
    3970                3975                3980
Asp Leu His Leu Arg Tyr Gln Lys Asp Lys Lys Gly Ile Ser Thr Ser
3985                3990                3995                4000
Ala Ala Ser Pro Ala Val Gly Thr Val Gly Met Asp Met Asp Glu Asp
                4005                4010                4015
Asp Asp Phe Ser Lys Trp Asn Phe Tyr Tyr Ser Pro Gln Ser Ser Pro
            4020                4025                4030
Asp Lys Lys Leu Thr Ile Phe Lys Thr Glu Leu Arg Val Arg Glu Ser
        4035                4040                4045
Asp Glu Glu Thr Gln Ile Lys Val Asn Trp Glu Glu Glu Ala Ala Ser
    4050                4055                4060
Gly Leu Leu Thr Ser Leu Lys Asp Asn Val Pro Lys Ala Thr Gly Val
4065                4070                4075                4080
Leu Tyr Asp Tyr Val Asn Lys Tyr His Trp Glu His Thr Gly Leu Thr
                4085                4090                4095
Leu Arg Glu Val Ser Ser Lys Leu Arg Arg Asn Leu Gln Asn Asn Ala
            4100                4105                4110
Glu Trp Val Tyr Gln Gly Ala Ile Arg Gln Ile Asp Asp Ile Asp Val
        4115                4120                4125
Arg Phe Gln Lys Ala Ala Ser Gly Thr Thr Gly Thr Tyr Gln Glu Trp
    4130                4135                4140
Lys Asp Lys Ala Gln Asn Leu Tyr Gln Glu Leu Leu Thr Gln Glu Gly
4145                4150                4155                4160
Gln Ala Ser Phe Gln Gly Leu Lys Asp Asn Val Phe Asp Gly Leu Val
                4165                4170                4175
Arg Val Thr Gln Lys Phe His Met Lys Val Lys His Leu Ile Asp Ser
            4180                4185                4190
Leu Ile Asp Phe Leu Asn Phe Pro Arg Phe Gln Phe Pro Gly Lys Pro
        4195                4200                4205
Gly Ile Tyr Thr Arg Glu Glu Leu Cys Thr Met Phe Ile Arg Glu Val
    4210                4215                4220
Gly Thr Val Leu Ser Gln Val Tyr Ser Lys Val His Asn Gly Ser Glu
4225                4230                4235                4240
Ile Leu Phe Ser Tyr Phe Gln Asp Leu Val Ile Thr Leu Pro Phe Glu
                4245                4250                4255
Leu Arg Lys His Lys Leu Ile Asp Val Ile Ser Met Tyr Arg Glu Leu
            4260                4265                4270
Leu Lys Asp Leu Ser Lys Glu Ala Gln Glu Val Phe Lys Ala Ile Gln
        4275                4280                4285
Ser Leu Lys Thr Thr Glu Val Leu Arg Asn Leu Gln Asp Leu Leu Gln
    4290                4295                4300
Phe Ile Phe Gln Leu Ile Glu Asp Asn Ile Lys Gln Leu Lys Glu Met
4305                4310                4315                4320
Lys Phe Thr Tyr Leu Ile Asn Tyr Ile Gln Asp Glu Ile Asn Thr Ile
                4325                4330                4335
Phe Asn Asp Tyr Ile Pro Tyr Val Phe Lys Leu Leu Lys Glu Asn Leu
            4340                4345                4350
Cys Leu Asn Leu His Lys Phe Asn Glu Phe Ile Gln Asn Glu Leu Gln
        4355                4360                4365
Glu Ala Ser Gln Glu Leu Gln Gln Ile His Gln Tyr Ile Met Ala Leu
    4370                4375                4380
```

```
Arg Glu Glu Tyr Phe Asp Pro Ser Ile Val Gly Trp Thr Val Lys Tyr
4385                4390                4395                4400
Tyr Glu Leu Glu Glu Lys Ile Val Ser Leu Ile Lys Asn Leu Leu Val
                4405                4410                4415
Ala Leu Lys Asp Phe His Ser Glu Tyr Ile Val Ser Ala Ser Asn Phe
            4420                4425                4430
Thr Ser Gln Leu Ser Ser Gln Val Glu Gln Phe Leu His Arg Asn Ile
        4435                4440                4445
Gln Glu Tyr Leu Ser Ile Leu Thr Asp Pro Asp Gly Lys Gly Lys Glu
    4450                4455                4460
Lys Ile Ala Glu Leu Ser Ala Thr Ala Gln Glu Ile Ile Lys Ser Gln
4465                4470                4475                4480
Ala Ile Ala Thr Lys Lys Ile Ile Ser Asp Tyr His Gln Gln Phe Arg
                4485                4490                4495
Tyr Lys Leu Gln Asp Phe Ser Asp Gln Leu Ser Asp Tyr Tyr Glu Lys
            4500                4505                4510
Phe Ile Ala Glu Ser Lys Arg Leu Ile Asp Leu Ser Ile Gln Asn Tyr
        4515                4520                4525
His Thr Phe Leu Ile Tyr Ile Thr Glu Leu Leu Lys Lys Leu Gln Ser
    4530                4535                4540
Thr Thr Val Met Asn Pro Tyr Met Lys Leu Ala Pro Gly Glu Leu Thr
4545                4550                4555                4560
Ile Ile Leu


<210> 33
<211> 2196
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (13)...(1983)
<223> Nucleotide sequence encoding
      5,10-methylenetetrahydofolate reductase (MTHFR)

<400> 33
aattccggag cc atg gtg aac gaa gcc aga gga aac agc agc ctc aac ccc      51
              Met Val Asn Glu Ala Arg Gly Asn Ser Ser Leu Asn Pro
               1               5                  10

tgc ttg gag ggc agt gcc agc agt ggc agt gag agc tcc aaa gat agt        99
Cys Leu Glu Gly Ser Ala Ser Ser Gly Ser Glu Ser Ser Lys Asp Ser
     15                  20                  25

tcg aga tgt tcc acc ccg ggc ctg gac cct gag cgg cat gag aga ctc       147
Ser Arg Cys Ser Thr Pro Gly Leu Asp Pro Glu Arg His Glu Arg Leu
 30                  35                  40                  45

cgg gag aag atg agg cgg cga ttg gaa tct ggt gac aag tgg ttc tcc       195
Arg Glu Lys Met Arg Arg Arg Leu Glu Ser Gly Asp Lys Trp Phe Ser
                 50                  55                  60

ctg gaa ttc ttc cct cct cga act gct gag gga gct gtc aat ctc atc       243
Leu Glu Phe Phe Pro Pro Arg Thr Ala Glu Gly Ala Val Asn Leu Ile
             65                  70                  75

tca agg ttt gac cgg atg gca gca ggt ggc ccc ctc tac ata gac gtg       291
```

Ser Arg Phe Asp Arg Met Ala Ala Gly Gly Pro Leu Tyr Ile Asp Val
80 85 90

acc tgg cac cca gca ggt gac cct ggc tca gac aag gag acc tcc tcc 339
Thr Trp His Pro Ala Gly Asp Pro Gly Ser Asp Lys Glu Thr Ser Ser
95 100 105

atg atg atc gcc agc acc gcc gtg aac tac tgt ggc ctg gag acc atc 387
Met Met Ile Ala Ser Thr Ala Val Asn Tyr Cys Gly Leu Glu Thr Ile
110 115 120 125

ctg cac atg acc tgc tgc cgt cag cgc ctg gag gag atc acg ggc cat 435
Leu His Met Thr Cys Cys Arg Gln Arg Leu Glu Glu Ile Thr Gly His
130 135 140

ctg cac aaa gct aag cag ctg ggc ctg aag aac atc atg gcg ctg cgg 483
Leu His Lys Ala Lys Gln Leu Gly Leu Lys Asn Ile Met Ala Leu Arg
145 150 155

gga gac cca ata ggt gac cag tgg gaa gag gag gag gga ggc ttc aac 531
Gly Asp Pro Ile Gly Asp Gln Trp Glu Glu Glu Glu Gly Gly Phe Asn
160 165 170

tac gca gtg gac ctg gtg aag cac atc cga agt gag ttt ggt gac tac 579
Tyr Ala Val Asp Leu Val Lys His Ile Arg Ser Glu Phe Gly Asp Tyr
175 180 185

ttt gac atc tgt gtg gca ggt tac ccc aaa ggc cac ccc gaa gca ggg 627
Phe Asp Ile Cys Val Ala Gly Tyr Pro Lys Gly His Pro Glu Ala Gly
190 195 200 205

agc ttt gag gct gac ctg aag cac ttg aag gag aag gtg tct gcg gga 675
Ser Phe Glu Ala Asp Leu Lys His Leu Lys Glu Lys Val Ser Ala Gly
210 215 220

gcc gat ttc atc atc acg cag ctt ttc ttt gag gct gac aca ttc ttc 723
Ala Asp Phe Ile Ile Thr Gln Leu Phe Phe Glu Ala Asp Thr Phe Phe
225 230 235

cgc ttt gtg aag gca tgc acc gac atg ggc atc act tgc ccc atc gtc 771
Arg Phe Val Lys Ala Cys Thr Asp Met Gly Ile Thr Cys Pro Ile Val
240 245 250

ccc ggg atc ttt ccc atc cag ggc tac cac tcc ctt cgg cag ctt gtg 819
Pro Gly Ile Phe Pro Ile Gln Gly Tyr His Ser Leu Arg Gln Leu Val
255 260 265

aag ctg tcc aag ctg gag gtg cca cag gag atc aag gac gtg att gag 867
Lys Leu Ser Lys Leu Glu Val Pro Gln Glu Ile Lys Asp Val Ile Glu
270 275 280 285

cca atc aaa gac aac gat gct gcc atc cgc aac tat ggc atc gag ctg 915
Pro Ile Lys Asp Asn Asp Ala Ala Ile Arg Asn Tyr Gly Ile Glu Leu
290 295 300

gcc gtg agc ctg tgc cag gag ctt ctg gcc agt ggc ttg gtg cca ggc 963
Ala Val Ser Leu Cys Gln Glu Leu Leu Ala Ser Gly Leu Val Pro Gly
305 310 315

```
ctc cac ttc tac acc ctc aac cgc gag atg gct acc aca gag gtg ctg      1011
Leu His Phe Tyr Thr Leu Asn Arg Glu Met Ala Thr Thr Glu Val Leu
        320                 325                 330

aag cgc ctg ggg atg tgg act gag gac ccc agg cgt ccc cta ccc tgg      1059
Lys Arg Leu Gly Met Trp Thr Glu Asp Pro Arg Arg Pro Leu Pro Trp
    335                 340                 345

gct ctc agt gcc cac ccc aag cgc cga gag gaa gat gta cgt ccc atc      1107
Ala Leu Ser Ala His Pro Lys Arg Arg Glu Glu Asp Val Arg Pro Ile
350                 355                 360                 365

ttc tgg gcc tcc aga cca aag agt tac atc tac cgt acc cag gag tgg      1155
Phe Trp Ala Ser Arg Pro Lys Ser Tyr Ile Tyr Arg Thr Gln Glu Trp
                370                 375                 380

gac gag ttc cct aac ggc cgc tgg ggc aat tcc tct tcc cct gcc ttt      1203
Asp Glu Phe Pro Asn Gly Arg Trp Gly Asn Ser Ser Ser Pro Ala Phe
            385                 390                 395

ggg gag ctg aag gac tac tac ctc ttc tac ctg aag agc aag tcc ccc      1251
Gly Glu Leu Lys Asp Tyr Tyr Leu Phe Tyr Leu Lys Ser Lys Ser Pro
        400                 405                 410

aag gag gag ctg ctg aag atg tgg ggg gag gag ctg acc agt gaa gca      1299
Lys Glu Glu Leu Leu Lys Met Trp Gly Glu Glu Leu Thr Ser Glu Ala
    415                 420                 425

agt gtc ttt gaa gtc ttt gtt ctt tac ctc tcg gga gaa cca aac cgg      1347
Ser Val Phe Glu Val Phe Val Leu Tyr Leu Ser Gly Glu Pro Asn Arg
430                 435                 440                 445

aat ggt cac aaa gtg act tgc ctg ccc tgg aac gat gag ccc ctg gcg      1395
Asn Gly His Lys Val Thr Cys Leu Pro Trp Asn Asp Glu Pro Leu Ala
                450                 455                 460

gct gag acc agc ctg ctg aag gag gag ctg ctg cgg gtg aac cgc cag      1443
Ala Glu Thr Ser Leu Leu Lys Glu Glu Leu Leu Arg Val Asn Arg Gln
            465                 470                 475

ggc atc ctc acc atc aac tca cag ccc aac atc aac ggg aag ccg tcc      1491
Gly Ile Leu Thr Ile Asn Ser Gln Pro Asn Ile Asn Gly Lys Pro Ser
        480                 485                 490

tcc gac ccc atc gtg ggc tgg ggc ccc agc ggg ggc tat gtc ttc cag      1539
Ser Asp Pro Ile Val Gly Trp Gly Pro Ser Gly Gly Tyr Val Phe Gln
    495                 500                 505

aag gcc tac tta gag ttt ttc act tcc cgc gag aca gcg gaa gca ctt      1587
Lys Ala Tyr Leu Glu Phe Phe Thr Ser Arg Glu Thr Ala Glu Ala Leu
510                 515                 520                 525

ctg caa gtg ctg aag aag tac gag ctc cgg gtt aat tac cac ctt gtc      1635
Leu Gln Val Leu Lys Lys Tyr Glu Leu Arg Val Asn Tyr His Leu Val
                530                 535                 540

aat gtg aag ggt gaa aac atc acc aat gcc cct gaa ctg cag ccg aat      1683
```

Asn Val Lys Gly Glu Asn Ile Thr Asn Ala Pro Glu Leu Gln Pro Asn
            545   550   555

gct gtc act tgg ggc atc ttc cct ggg cga gag atc atc cag ccc acc   1731
Ala Val Thr Trp Gly Ile Phe Pro Gly Arg Glu Ile Ile Gln Pro Thr
        560   565   570

gta gtg gat ccc gtc agc ttc atg ttc tgg aag gac gag gcc ttt gcc   1779
Val Val Asp Pro Val Ser Phe Met Phe Trp Lys Asp Glu Ala Phe Ala
    575   580   585

ctg tgg att gag cgg tgg gga aag ctg tat gag gag gag tcc ccg tcc   1827
Leu Trp Ile Glu Arg Trp Gly Lys Leu Tyr Glu Glu Glu Ser Pro Ser
590   595   600   605

cgc acc atc atc cag tac atc cac gac aac tac ttc ctg gtc aac ctg   1875
Arg Thr Ile Ile Gln Tyr Ile His Asp Asn Tyr Phe Leu Val Asn Leu
                610   615   620

gtg gac aat gac ttc cca ctg gac aac tgc ctc tgg cag gtg gtg gaa   1923
Val Asp Asn Asp Phe Pro Leu Asp Asn Cys Leu Trp Gln Val Val Glu
            625   630   635

gac aca ttg gag ctt ctc aac agg ccc acc cag aat gcg aga gaa acg   1971
Asp Thr Leu Glu Leu Leu Asn Arg Pro Thr Gln Asn Ala Arg Glu Thr
        640   645   650

gag gct cca tga ccctgcgtcc tgacgccctg cgttggagcc actcctgtcc   2023
Glu Ala Pro *
    655

cgccttcctc ctccacagtg ctgcttctct tgggaactcc actctccttc gtgtctctcc   2083
caccccggcc tccactcccc cacctgacaa tggcagctag actggagtga ggcttccagg   2143
ctcttcctgg acctgagtcg gccccacatg ggaacctagt actctctgct cta   2196

<210> 34
<211> 656
<212> PRT
<213> Homo sapien

<400> 34
Met Val Asn Glu Ala Arg Gly Asn Ser Ser Leu Asn Pro Cys Leu Glu
1   5   10   15
Gly Ser Ala Ser Ser Gly Ser Glu Ser Ser Lys Asp Ser Ser Arg Cys
            20   25   30
Ser Thr Pro Gly Leu Asp Pro Glu Arg His Glu Arg Leu Arg Glu Lys
        35   40   45
Met Arg Arg Arg Leu Glu Ser Gly Asp Lys Trp Phe Ser Leu Glu Phe
    50   55   60
Phe Pro Pro Arg Thr Ala Glu Gly Ala Val Asn Leu Ile Ser Arg Phe
65   70   75   80
Asp Arg Met Ala Ala Gly Gly Pro Leu Tyr Ile Asp Val Thr Trp His
                85   90   95
Pro Ala Gly Asp Pro Gly Ser Asp Lys Glu Thr Ser Ser Met Met Ile
            100   105   110
Ala Ser Thr Ala Val Asn Tyr Cys Gly Leu Glu Thr Ile Leu His Met
        115   120   125
Thr Cys Cys Arg Gln Arg Leu Glu Glu Ile Thr Gly His Leu His Lys

    130   135   140
Ala Lys Gln Leu Gly Leu Lys Asn Ile Met Ala Leu Arg Gly Asp Pro
145   150   155   160
Ile Gly Asp Gln Trp Glu Glu Glu Glu Gly Gly Phe Asn Tyr Ala Val
                165   170   175
Asp Leu Val Lys His Ile Arg Ser Glu Phe Gly Asp Tyr Phe Asp Ile
            180   185   190
Cys Val Ala Gly Tyr Pro Lys Gly His Pro Glu Ala Gly Ser Phe Glu
        195   200   205
Ala Asp Leu Lys His Leu Lys Glu Lys Val Ser Ala Gly Ala Asp Phe
    210   215   220
Ile Ile Thr Gln Leu Phe Phe Glu Ala Asp Thr Phe Phe Arg Phe Val
225   230   235   240
Lys Ala Cys Thr Asp Met Gly Ile Thr Cys Pro Ile Val Pro Gly Ile
                245   250   255
Phe Pro Ile Gln Gly Tyr His Ser Leu Arg Gln Leu Val Lys Leu Ser
            260   265   270
Lys Leu Glu Val Pro Gln Glu Ile Lys Asp Val Ile Glu Pro Ile Lys
        275   280   285
Asp Asn Asp Ala Ala Ile Arg Asn Tyr Gly Ile Glu Leu Ala Val Ser
    290   295   300
Leu Cys Gln Glu Leu Leu Ala Ser Gly Leu Val Pro Gly Leu His Phe
305   310   315   320
Tyr Thr Leu Asn Arg Glu Met Ala Thr Thr Glu Val Leu Lys Arg Leu
                325   330   335
Gly Met Trp Thr Glu Asp Pro Arg Arg Pro Leu Pro Trp Ala Leu Ser
            340   345   350
Ala His Pro Lys Arg Arg Glu Glu Asp Val Arg Pro Ile Phe Trp Ala
        355   360   365
Ser Arg Pro Lys Ser Tyr Ile Tyr Arg Thr Gln Glu Trp Asp Glu Phe
    370   375   380
Pro Asn Gly Arg Trp Gly Asn Ser Ser Ser Pro Ala Phe Gly Glu Leu
385   390   395   400
Lys Asp Tyr Tyr Leu Phe Tyr Leu Lys Ser Lys Ser Pro Lys Glu Glu
                405   410   415
Leu Leu Lys Met Trp Gly Glu Glu Leu Thr Ser Glu Ala Ser Val Phe
            420   425   430
Glu Val Phe Val Leu Tyr Leu Ser Gly Glu Pro Asn Arg Asn Gly His
        435   440   445
Lys Val Thr Cys Leu Pro Trp Asn Asp Glu Pro Leu Ala Ala Glu Thr
    450   455   460
Ser Leu Leu Lys Glu Glu Leu Leu Arg Val Asn Arg Gln Gly Ile Leu
465   470   475   480
Thr Ile Asn Ser Gln Pro Asn Ile Asn Gly Lys Pro Ser Ser Asp Pro
                485   490   495
Ile Val Gly Trp Gly Pro Ser Gly Gly Tyr Val Phe Gln Lys Ala Tyr
            500   505   510
Leu Glu Phe Phe Thr Ser Arg Glu Thr Ala Glu Ala Leu Leu Gln Val
        515   520   525
Leu Lys Lys Tyr Glu Leu Arg Val Asn Tyr His Leu Val Asn Val Lys
    530   535   540
Gly Glu Asn Ile Thr Asn Ala Pro Glu Leu Gln Pro Asn Ala Val Thr
545   550   555   560
Trp Gly Ile Phe Pro Gly Arg Glu Ile Ile Gln Pro Thr Val Val Asp
                565   570   575
Pro Val Ser Phe Met Phe Trp Lys Asp Glu Ala Phe Ala Leu Trp Ile
            580   585   590
Glu Arg Trp Gly Lys Leu Tyr Glu Glu Glu Ser Pro Ser Arg Thr Ile

```
        595                 600                 605
Ile Gln Tyr Ile His Asp Asn Tyr Phe Leu Val Asn Leu Val Asp Asn
    610                 615                 620
Asp Phe Pro Leu Asp Asn Cys Leu Trp Gln Val Val Glu Asp Thr Leu
625                 630                 635                 640
Glu Leu Leu Asn Arg Pro Thr Gln Asn Ala Arg Glu Thr Glu Ala Pro
                645                 650                 655

<210> 35
<211> 3834
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (117)...(1949)
<223> Nucleotide sequence encoding selectin E (SELE)

<400> 35
cctgagacag aggcagcagt gatacccacc tgagagatcc tgtgtttgaa caactgcttc     60
ccaaaacgga aagtatttca agcctaaacc tttgggtgaa aagaactctt gaagtc atg    119
                                                              Met
                                                               1

att gct tca cag ttt ctc tca gct ctc act ttg gtg ctt ctc att aaa      167
Ile Ala Ser Gln Phe Leu Ser Ala Leu Thr Leu Val Leu Leu Ile Lys
            5                  10                  15

gag agt gga gcc tgg tct tac aac acc tcc acg gaa gct atg act tat      215
Glu Ser Gly Ala Trp Ser Tyr Asn Thr Ser Thr Glu Ala Met Thr Tyr
        20                  25                  30

gat gag gcc agt gct tat tgt cag caa agg tac aca cac ctg gtt gca      263
Asp Glu Ala Ser Ala Tyr Cys Gln Gln Arg Tyr Thr His Leu Val Ala
    35                  40                  45

att caa aac aaa gaa gag att gag tac cta aac tcc ata ttg agc tat      311
Ile Gln Asn Lys Glu Glu Ile Glu Tyr Leu Asn Ser Ile Leu Ser Tyr
50                  55                  60                  65

tca cca agt tat tac tgg att gga atc aga aaa gtc aac aat gtg tgg      359
Ser Pro Ser Tyr Tyr Trp Ile Gly Ile Arg Lys Val Asn Asn Val Trp
                70                  75                  80

gtc tgg gta gga acc cag aaa cct ctg aca gaa gaa gcc aag aac tgg      407
Val Trp Val Gly Thr Gln Lys Pro Leu Thr Glu Glu Ala Lys Asn Trp
            85                  90                  95

gct cca ggt gaa ccc aac aat agg caa aaa gat gag gac tgc gtg gag      455
Ala Pro Gly Glu Pro Asn Asn Arg Gln Lys Asp Glu Asp Cys Val Glu
        100                 105                 110

atc tac atc aag aga gaa aaa gat gtg ggc atg tgg aat gat gag agg      503
Ile Tyr Ile Lys Arg Glu Lys Asp Val Gly Met Trp Asn Asp Glu Arg
    115                 120                 125

tgc agc aag aag aag ctt gcc cta tgc tac aca gct gcc tgt acc aat      551
Cys Ser Lys Lys Lys Leu Ala Leu Cys Tyr Thr Ala Ala Cys Thr Asn
```

```
130                 135                 140                 145

aca tcc tgc agt ggc cac ggt gaa tgt gta gag acc atc aat aat tac      599
Thr Ser Cys Ser Gly His Gly Glu Cys Val Glu Thr Ile Asn Asn Tyr
                150                 155                 160

act tgc aag tgt gac cct ggc ttc agt gga ctc aag tgt gag caa att      647
Thr Cys Lys Cys Asp Pro Gly Phe Ser Gly Leu Lys Cys Glu Gln Ile
            165                 170                 175

gtg aac tgt aca gcc ctg gaa tcc cct gag cat gga agc ctg gtt tgc      695
Val Asn Cys Thr Ala Leu Glu Ser Pro Glu His Gly Ser Leu Val Cys
        180                 185                 190

agt cac cca ctg gga aac ttc agc tac aat tct tcc tgc tct atc agc      743
Ser His Pro Leu Gly Asn Phe Ser Tyr Asn Ser Ser Cys Ser Ile Ser
    195                 200                 205

tgt gat agg ggt tac ctg cca agc agc atg gag acc atg cag tgt atg      791
Cys Asp Arg Gly Tyr Leu Pro Ser Ser Met Glu Thr Met Gln Cys Met
210                 215                 220                 225

tcc tct gga gaa tgg agt gct cct att cca gcc tgc aat gtg gtt gag      839
Ser Ser Gly Glu Trp Ser Ala Pro Ile Pro Ala Cys Asn Val Val Glu
                230                 235                 240

tgt gat gct gtg aca aat cca gcc aat ggg ttc gtg gaa tgt ttc caa      887
Cys Asp Ala Val Thr Asn Pro Ala Asn Gly Phe Val Glu Cys Phe Gln
            245                 250                 255

aac cct gga agc ttc cca tgg aac aca acc tgt aca ttt gac tgt gaa      935
Asn Pro Gly Ser Phe Pro Trp Asn Thr Thr Cys Thr Phe Asp Cys Glu
        260                 265                 270

gaa gga ttt gaa cta atg gga gcc cag agc ctt cag tgt acc tca tct      983
Glu Gly Phe Glu Leu Met Gly Ala Gln Ser Leu Gln Cys Thr Ser Ser
    275                 280                 285

ggg aat tgg gac aac gag aag cca acg tgt aaa gct gtg aca tgc agg     1031
Gly Asn Trp Asp Asn Glu Lys Pro Thr Cys Lys Ala Val Thr Cys Arg
290                 295                 300                 305

gcc gtc cgc cag cct cag aat ggc tct gtg agg tgc agc cat tcc cct     1079
Ala Val Arg Gln Pro Gln Asn Gly Ser Val Arg Cys Ser His Ser Pro
                310                 315                 320

gct gga gag ttc acc ttc aaa tca tcc tgc aac ttc acc tgt gag gaa     1127
Ala Gly Glu Phe Thr Phe Lys Ser Ser Cys Asn Phe Thr Cys Glu Glu
            325                 330                 335

ggc ttc atg ttg cag gga cca gcc cag gtt gaa tgc acc act caa ggg     1175
Gly Phe Met Leu Gln Gly Pro Ala Gln Val Glu Cys Thr Thr Gln Gly
        340                 345                 350

cag tgg aca cag caa atc cca gtt tgt gaa gct ttc cag tgc aca gcc     1223
Gln Trp Thr Gln Gln Ile Pro Val Cys Glu Ala Phe Gln Cys Thr Ala
    355                 360                 365
```

```
ttg tcc aac ccc gag cga ggc tac atg aat tgt ctt cct agt gct tct      1271
Leu Ser Asn Pro Glu Arg Gly Tyr Met Asn Cys Leu Pro Ser Ala Ser
370                 375                 380                 385

ggc agt ttc cgt tat ggg tcc agc tgt gag ttc tcc tgt gag cag ggt      1319
Gly Ser Phe Arg Tyr Gly Ser Ser Cys Glu Phe Ser Cys Glu Gln Gly
                390                 395                 400

ttt gtg ttg aag gga tcc aaa agg ctc caa tgt ggc ccc aca ggg gag      1367
Phe Val Leu Lys Gly Ser Lys Arg Leu Gln Cys Gly Pro Thr Gly Glu
            405                 410                 415

tgg gac aac gag aag ccc aca tgt gaa gct gtg aga tgc gat gct gtc      1415
Trp Asp Asn Glu Lys Pro Thr Cys Glu Ala Val Arg Cys Asp Ala Val
        420                 425                 430

cac cag ccc ccg aag ggt ttg gtg agg tgt gct cat tcc cct att gga      1463
His Gln Pro Pro Lys Gly Leu Val Arg Cys Ala His Ser Pro Ile Gly
    435                 440                 445

gaa ttc acc tac aag tcc tct tgt gcc ttc agc tgt gag gag gga ttt      1511
Glu Phe Thr Tyr Lys Ser Ser Cys Ala Phe Ser Cys Glu Glu Gly Phe
450                 455                 460                 465

gaa tta tat gga tca act caa ctt gag tgc aca tct cag gga caa tgg      1559
Glu Leu Tyr Gly Ser Thr Gln Leu Glu Cys Thr Ser Gln Gly Gln Trp
                470                 475                 480

aca gaa gag gtt cct tcc tgc caa gtg gta aaa tgt tca agc ctg gca      1607
Thr Glu Glu Val Pro Ser Cys Gln Val Val Lys Cys Ser Ser Leu Ala
            485                 490                 495

gtt ccg gga aag atc aac atg agc tgc agt ggg gag ccc gtg ttt ggc      1655
Val Pro Gly Lys Ile Asn Met Ser Cys Ser Gly Glu Pro Val Phe Gly
        500                 505                 510

act gtg tgc aag ttc gcc tgt cct gaa gga tgg acg ctc aat ggc tct      1703
Thr Val Cys Lys Phe Ala Cys Pro Glu Gly Trp Thr Leu Asn Gly Ser
    515                 520                 525

gca gct cgg aca tgt gga gcc aca gga cac tgg tct ggc ctg cta cct      1751
Ala Ala Arg Thr Cys Gly Ala Thr Gly His Trp Ser Gly Leu Leu Pro
530                 535                 540                 545

acc tgt gaa gct ccc act gag tcc aac att ccc ttg gta gct gga ctt      1799
Thr Cys Glu Ala Pro Thr Glu Ser Asn Ile Pro Leu Val Ala Gly Leu
                550                 555                 560

tct gct gct gga ctc tcc ctc ctg aca tta gca cca ttt ctc ctc tgg      1847
Ser Ala Ala Gly Leu Ser Leu Leu Thr Leu Ala Pro Phe Leu Leu Trp
            565                 570                 575

ctt cgg aaa tgc tta cgg aaa gca aag aaa ttt gtt cct gcc agc agc      1895
Leu Arg Lys Cys Leu Arg Lys Ala Lys Lys Phe Val Pro Ala Ser Ser
        580                 585                 590

tgc caa agc ctt gaa tca gac gga agc tac caa aag cct tct tac atc      1943
Cys Gln Ser Leu Glu Ser Asp Gly Ser Tyr Gln Lys Pro Ser Tyr Ile
```

```
    595                 600                 605

ctt taa gttcaaaaga atcagaaaca ggtgcatctg gggaactaga gggatacact      1999
Leu  *
610

gaagttaaca gagacagata actctcctcg ggtctctggc ccttcttgcc tactatgcca   2059
gatgccttta tggctgaaac cgcaacaccc atcaccactt caatagatca aagtccagca   2119
ggcaaggacg gccttcaact gaaaagactc agtgttccct ttcctactct caggatcaag   2179
aaagtgttgg ctaatgaagg gaaaggatat tttcttccaa gcaaaggtga agagaccaag   2239
actctgaaat ctcagaattc cttttctaac tctcccttgc tcgctgtaaa atcttggcac   2299
agaaacacaa tattttgtgg ctttctttct tttgcccttc acagtgtttc gacagctgat   2359
tacacagttg ctgtcataag aatgaataat aattatccag agtttagagg aaaaaaatga   2419
ctaaaaatat tataacttaa aaaaatgaca gatgttgaat gcccacaggc aaatgcatgg   2479
agggttgtta atggtgcaaa tcctactgaa tgctctgtgc gagggttact atgcacaatt   2539
taatcacttt catccctatg ggattcagtg cttcttaaag agttcttaag gattgtgata   2599
tttttacttg cattgaatat attataatct tccatacttc ttcattcaat acaagtgtgg   2659
tagggactta aaaaacttgt aaatgctgtc aactatgata tggtaaaagt tacttattct   2719
agattacccc ctcattgttt attaacaaat tatgttacat ctgttttaaa tttatttcaa   2779
aaagggaaac tattgtcccc tagcaaggca tgatgttaac cagaataaag ttctgagtgt   2839
ttttactaca gttgtttttt gaaaacatgg tagaattgga gagtaaaaac tgaatggaag   2899
gtttgtatat tgtcagatat tttttcagaa atatgtggtt tccacgatga aaaacttcca   2959
tgaggccaaa cgttttgaac taataaaagc ataaatgcaa acacacaaag gtataatttt   3019
atgaatgtct ttgttggaaa agaatacaga aagatggatg tgctttgcat tcctacaaag   3079
atgtttgtca gatgtgatat gtaaacataa ttcttgtata ttatggaaga ttttaaattc   3139
acaatagaaa ctcaccatgt aaaagagtca tctggtagat tttaacgaa tgaagatgtc    3199
taatagttat tccctatttg ttttcttctg tatgttaggg tgctctggaa gagaggaatg   3259
cctgtgtgag caagcattta tgtttattta taagcagatt taacaattcc aaaggaatct   3319
ccagttttca gttgatcact ggcaatgaaa aattctcagt cagtaattgc caaagctgct   3379
ctagccttga ggagtgtgag aatcaaaact ctcctacact tccattaact tagcatgtgt   3439
tgaaaaaaaa agtttcagag aagttctggc tgaacactgg caacgacaaa gccaacagtc   3499
aaaacagaga tgtgataagg atcagaacag cagaggttct tttaaagggg cagaaaaact   3559
ctgggaaata agagagaaca actactgtga tcaggctatg tatggaatac agtgttattt   3619
tctttgaaat tgtttaagtg ttgtaaatat ttatgtaaac tgcattagaa attagctgtg   3679
tgaaatacca gtgtggtttg tgtttgagtt ttattgagaa ttttaaatta taacttaaaa   3739
tattttataa tttttaaagt atatatttat ttaagcttat gtcagaccta tttgacataa   3799
cactataaag gttgacaata aatgtgctta tgttt                              3834

<210> 36
<211> 610
<212> PRT
<213> Homo sapien

<400> 36
Met Ile Ala Ser Gln Phe Leu Ser Ala Leu Thr Leu Val Leu Leu Ile
 1               5                  10                  15
Lys Glu Ser Gly Ala Trp Ser Tyr Asn Thr Ser Thr Glu Ala Met Thr
            20                  25                  30
Tyr Asp Glu Ala Ser Ala Tyr Cys Gln Gln Arg Tyr Thr His Leu Val
        35                  40                  45
Ala Ile Gln Asn Lys Glu Glu Ile Glu Tyr Leu Asn Ser Ile Leu Ser
    50                  55                  60
Tyr Ser Pro Ser Tyr Tyr Trp Ile Gly Ile Arg Lys Val Asn Asn Val
65                  70                  75                  80
Trp Val Trp Val Gly Thr Gln Lys Pro Leu Thr Glu Glu Ala Lys Asn
                85                  90                  95
Trp Ala Pro Gly Glu Pro Asn Asn Arg Gln Lys Asp Glu Asp Cys Val
            100                 105                 110
```

```
Glu Ile Tyr Ile Lys Arg Glu Lys Asp Val Gly Met Trp Asn Asp Glu
        115                 120                 125
Arg Cys Ser Lys Lys Lys Leu Ala Leu Cys Tyr Thr Ala Ala Cys Thr
    130                 135                 140
Asn Thr Ser Cys Ser Gly His Gly Glu Cys Val Glu Thr Ile Asn Asn
145                 150                 155                 160
Tyr Thr Cys Lys Cys Asp Pro Gly Phe Ser Gly Leu Lys Cys Glu Gln
                165                 170                 175
Ile Val Asn Cys Thr Ala Leu Glu Ser Pro Glu His Gly Ser Leu Val
            180                 185                 190
Cys Ser His Pro Leu Gly Asn Phe Ser Tyr Asn Ser Ser Cys Ser Ile
        195                 200                 205
Ser Cys Asp Arg Gly Tyr Leu Pro Ser Ser Met Glu Thr Met Gln Cys
    210                 215                 220
Met Ser Ser Gly Glu Trp Ser Ala Pro Ile Pro Ala Cys Asn Val Val
225                 230                 235                 240
Glu Cys Asp Ala Val Thr Asn Pro Ala Asn Gly Phe Val Glu Cys Phe
                245                 250                 255
Gln Asn Pro Gly Ser Phe Pro Trp Asn Thr Thr Cys Thr Phe Asp Cys
            260                 265                 270
Glu Glu Gly Phe Glu Leu Met Gly Ala Gln Ser Leu Gln Cys Thr Ser
        275                 280                 285
Ser Gly Asn Trp Asp Asn Glu Lys Pro Thr Cys Lys Ala Val Thr Cys
    290                 295                 300
Arg Ala Val Arg Gln Pro Gln Asn Gly Ser Val Arg Cys Ser His Ser
305                 310                 315                 320
Pro Ala Gly Glu Phe Thr Phe Lys Ser Ser Cys Asn Phe Thr Cys Glu
                325                 330                 335
Glu Gly Phe Met Leu Gln Gly Pro Ala Gln Val Glu Cys Thr Thr Gln
            340                 345                 350
Gly Gln Trp Thr Gln Gln Ile Pro Val Cys Glu Ala Phe Gln Cys Thr
        355                 360                 365
Ala Leu Ser Asn Pro Glu Arg Gly Tyr Met Asn Cys Leu Pro Ser Ala
    370                 375                 380
Ser Gly Ser Phe Arg Tyr Gly Ser Ser Cys Glu Phe Ser Cys Glu Gln
385                 390                 395                 400
Gly Phe Val Leu Lys Gly Ser Lys Arg Leu Gln Cys Gly Pro Thr Gly
                405                 410                 415
Glu Trp Asp Asn Glu Lys Pro Thr Cys Glu Ala Val Arg Cys Asp Ala
            420                 425                 430
Val His Gln Pro Pro Lys Gly Leu Val Arg Cys Ala His Ser Pro Ile
        435                 440                 445
Gly Glu Phe Thr Tyr Lys Ser Ser Cys Ala Phe Ser Cys Glu Glu Gly
    450                 455                 460
Phe Glu Leu Tyr Gly Ser Thr Gln Leu Glu Cys Thr Ser Gln Gly Gln
465                 470                 475                 480
Trp Thr Glu Glu Val Pro Ser Cys Gln Val Val Lys Cys Ser Ser Leu
                485                 490                 495
Ala Val Pro Gly Lys Ile Asn Met Ser Cys Ser Gly Glu Pro Val Phe
            500                 505                 510
Gly Thr Val Cys Lys Phe Ala Cys Pro Glu Gly Trp Thr Leu Asn Gly
        515                 520                 525
Ser Ala Ala Arg Thr Cys Gly Ala Thr Gly His Trp Ser Gly Leu Leu
    530                 535                 540
Pro Thr Cys Glu Ala Pro Thr Glu Ser Asn Ile Pro Leu Val Ala Gly
545                 550                 555                 560
Leu Ser Ala Ala Gly Leu Ser Leu Leu Thr Leu Ala Pro Phe Leu Leu
                565                 570                 575
```

```
Trp Leu Arg Lys Cys Leu Arg Lys Ala Lys Lys Phe Val Pro Ala Ser
            580                 585                 590
Ser Cys Gln Ser Leu Glu Ser Asp Gly Ser Tyr Gln Lys Pro Ser Tyr
        595                 600                 605
Ile Leu
    610

<210> 37
<211> 1922
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (406)...(1428)
<223> Nucleotide sequence encoding nucleotide binding
      protein (G Protein), beta polypeptide 3 (GNB3)

<400> 37
ccacaatagg ggcagacctg tccatccttc tctgtgggtc ccctgtacct ttctccccca      60
acaggatcag acccagaggc agctggttgg ggtttgtcga gaagaaggat tatccagatc     120
agtcctttct aatctcagct cctgcctgta ccctcccata ctcaccaaac cctcttcccc     180
accaccctga gctgaggagc acagtttgag gcccccccaa ccccccgccg gtcggggcca     240
ggccaggcca ggccagctcc tctggcagca gagcctgggc aggtgacggg cgggcgcggg     300
cgtcgcagct gagggagtaa ggaggctccc aggaaccgga gctggaaacc cggccgaggt     360
ccagccagag cccaagagcc agagtgaccc ctcgacctgt cagcc atg ggg gag atg     417
                                                  Met Gly Glu Met
                                                   1

gag caa ctg cgt cag gaa gcg gag cag ctc aag aag cag att gca gat       465
Glu Gln Leu Arg Gln Glu Ala Glu Gln Leu Lys Lys Gln Ile Ala Asp
 5                  10                  15                  20

gcc agg aaa gcc tgt gct gac gtt act ctg gca gag ctg gtg tct ggc       513
Ala Arg Lys Ala Cys Ala Asp Val Thr Leu Ala Glu Leu Val Ser Gly
                25                  30                  35

cta gag gtg gtg gga cga gtc cag atg cgg acg cgg cgg acg tta agg       561
Leu Glu Val Val Gly Arg Val Gln Met Arg Thr Arg Arg Thr Leu Arg
            40                  45                  50

gga cac ctg gcc aag att tac gcc atg cac tgg gcc act gat tct aag       609
Gly His Leu Ala Lys Ile Tyr Ala Met His Trp Ala Thr Asp Ser Lys
        55                  60                  65

ctg ctg gta agt gcc tcg caa gat ggg aag ctg atc gtg tgg gac agc       657
Leu Leu Val Ser Ala Ser Gln Asp Gly Lys Leu Ile Val Trp Asp Ser
    70                  75                  80

tac acc acc aac aag gtg cac gcc atc cca ctg cgc tcc tcc tgg gtc       705
Tyr Thr Thr Asn Lys Val His Ala Ile Pro Leu Arg Ser Ser Trp Val
 85                  90                  95                 100

atg acc tgt gcc tat gcc cca tca ggg aac ttt gtg gca tgt ggg ggg       753
Met Thr Cys Ala Tyr Ala Pro Ser Gly Asn Phe Val Ala Cys Gly Gly
                105                 110                 115

ctg gac aac atg tgt tcc atc tac aac ctc aaa tcc cgt gag ggc aat       801
```

```
Leu Asp Asn Met Cys Ser Ile Tyr Asn Leu Lys Ser Arg Glu Gly Asn
            120                 125                 130

gtc aag gtc agc cgg gag ctt tct gct cac aca ggt tat ctc tcc tgc       849
Val Lys Val Ser Arg Glu Leu Ser Ala His Thr Gly Tyr Leu Ser Cys
        135                 140                 145

tgc cgc ttc ctg gat gac aac aat att gtg acc agc tcg ggg gac acc       897
Cys Arg Phe Leu Asp Asp Asn Asn Ile Val Thr Ser Ser Gly Asp Thr
    150                 155                 160

acg tgt gcc ttg tgg gac att gag act ggg cag cag aag act gta ttt       945
Thr Cys Ala Leu Trp Asp Ile Glu Thr Gly Gln Gln Lys Thr Val Phe
165                 170                 175                 180

gtg gga cac acg ggt gac tgc atg agc ctg gct gtg tct cct gac ttc       993
Val Gly His Thr Gly Asp Cys Met Ser Leu Ala Val Ser Pro Asp Phe
                185                 190                 195

aat ctc ttc att tcg ggg gcc tgt gat gcc agt gcc aag ctc tgg gat      1041
Asn Leu Phe Ile Ser Gly Ala Cys Asp Ala Ser Ala Lys Leu Trp Asp
            200                 205                 210

gtg cga gag ggg acc tgc cgt cag act ttc act ggc cac gag tcg gac      1089
Val Arg Glu Gly Thr Cys Arg Gln Thr Phe Thr Gly His Glu Ser Asp
        215                 220                 225

atc aac gcc atc tgt ttc ttc ccc aat gga gag gcc atc tgc acg ggc      1137
Ile Asn Ala Ile Cys Phe Phe Pro Asn Gly Glu Ala Ile Cys Thr Gly
    230                 235                 240

tcg gat gac gct tcc tgc cgc ttg ttt gac ctg cgg gca gac cag gag      1185
Ser Asp Asp Ala Ser Cys Arg Leu Phe Asp Leu Arg Ala Asp Gln Glu
245                 250                 255                 260

ctg atc tgc ttc tcc cac gag agc atc atc tgc ggc atc acg tcc gtg      1233
Leu Ile Cys Phe Ser His Glu Ser Ile Ile Cys Gly Ile Thr Ser Val
                265                 270                 275

gcc ttc tcc ctc agt ggc cgc cta cta ttc gct ggc tac gac gac ttc      1281
Ala Phe Ser Leu Ser Gly Arg Leu Leu Phe Ala Gly Tyr Asp Asp Phe
            280                 285                 290

aac tgc aat gtc tgg gac tcc atg aag tct gag cgt gtg ggc atc ctc      1329
Asn Cys Asn Val Trp Asp Ser Met Lys Ser Glu Arg Val Gly Ile Leu
        295                 300                 305

tct ggc cac gat aac agg gtg agc tgc ctg gga gtc aca gct gac ggg      1377
Ser Gly His Asp Asn Arg Val Ser Cys Leu Gly Val Thr Ala Asp Gly
    310                 315                 320

atg gct gtg gcc aca ggt tcc tgg gac agc ttc ctc aaa atc tgg aac      1425
Met Ala Val Ala Thr Gly Ser Trp Asp Ser Phe Leu Lys Ile Trp Asn
325                 330                 335                 340

tga ggaggctgga gaaagggaag tggaaggcag tgaacacact cagcagcccc           1478
 *
```

```
ctgcccgacc ccatctcatt caggtgttct cttctatatt ccgggtgcca ttcccactaa    1538
gctttctcct ttgagggcag tggggagcat gggactgtgc ctttgggagg cagcatcagg    1598
gacacagggg caaagaactg ccccatctcc tcccatggcc ttccctcccc acagtcctca    1658
cagcctctcc cttaatgagc aaggacaacc tgcccctccc cagccctttg caggcccagc    1718
agacttgagt ctgaggcccc aggccctagg attcctcccc cagagccact acctttgtcc    1778
aggcctgggt ggtatagggc gtttggccct gtgactatgg ctctggcacc actagggtcc    1838
tggccctctt cttattcatg ctttctcctt tttctacctt tttttctctc ctaagacacc    1898
tgcaataaag tgtagcaccc tggt                                            1922

<210> 38
<211> 340
<212> PRT
<213> Homo sapien

<400> 38
Met Gly Glu Met Glu Gln Leu Arg Gln Glu Ala Glu Gln Leu Lys Lys
 1               5                  10                  15
Gln Ile Ala Asp Ala Arg Lys Ala Cys Ala Asp Val Thr Leu Ala Glu
            20                  25                  30
Leu Val Ser Gly Leu Glu Val Val Gly Arg Val Gln Met Arg Thr Arg
        35                  40                  45
Arg Thr Leu Arg Gly His Leu Ala Lys Ile Tyr Ala Met His Trp Ala
    50                  55                  60
Thr Asp Ser Lys Leu Leu Val Ser Ala Ser Gln Asp Gly Lys Leu Ile
65                  70                  75                  80
Val Trp Asp Ser Tyr Thr Thr Asn Lys Val His Ala Ile Pro Leu Arg
                85                  90                  95
Ser Ser Trp Val Met Thr Cys Ala Tyr Ala Pro Ser Gly Asn Phe Val
            100                 105                 110
Ala Cys Gly Gly Leu Asp Asn Met Cys Ser Ile Tyr Asn Leu Lys Ser
        115                 120                 125
Arg Glu Gly Asn Val Lys Val Ser Arg Glu Leu Ser Ala His Thr Gly
    130                 135                 140
Tyr Leu Ser Cys Cys Arg Phe Leu Asp Asp Asn Asn Ile Val Thr Ser
145                 150                 155                 160
Ser Gly Asp Thr Thr Cys Ala Leu Trp Asp Ile Glu Thr Gly Gln Gln
                165                 170                 175
Lys Thr Val Phe Val Gly His Thr Gly Asp Cys Met Ser Leu Ala Val
            180                 185                 190
Ser Pro Asp Phe Asn Leu Phe Ile Ser Gly Ala Cys Asp Ala Ser Ala
        195                 200                 205
Lys Leu Trp Asp Val Arg Glu Gly Thr Cys Arg Gln Thr Phe Thr Gly
    210                 215                 220
His Glu Ser Asp Ile Asn Ala Ile Cys Phe Phe Pro Asn Gly Glu Ala
225                 230                 235                 240
Ile Cys Thr Gly Ser Asp Asp Ala Ser Cys Arg Leu Phe Asp Leu Arg
                245                 250                 255
Ala Asp Gln Glu Leu Ile Cys Phe Ser His Glu Ser Ile Ile Cys Gly
            260                 265                 270
Ile Thr Ser Val Ala Phe Ser Leu Ser Gly Arg Leu Leu Phe Ala Gly
        275                 280                 285
Tyr Asp Asp Phe Asn Cys Asn Val Trp Asp Ser Met Lys Ser Glu Arg
    290                 295                 300
Val Gly Ile Leu Ser Gly His Asp Asn Arg Val Ser Cys Leu Gly Val
305                 310                 315                 320
Thr Ala Asp Gly Met Ala Val Ala Thr Gly Ser Trp Asp Ser Phe Leu
                325                 330                 335
```

Lys Ile Trp Asn
340

<210> 39
<211> 2443
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (162)...(1253)
<223> Nucleotide sequence encoding angiotensin receptor
2 (AGTR2)

<400> 39
acgtcccagc gtctgagaga acgagtaagc aagaattcaa agcattctgc agcctgaatt 60
ttgaaggagt gtgtttaggc actaagcaag ctgatttatg ataactgctt taaacttcaa 120
caaccaaagg cataagaact aggagctgct gacatttcaa t atg aag ggc aac tcc 176
Met Lys Gly Asn Ser
1 5

acc ctt gcc act act agc aaa aac att acc agc ggt ctt cac ttc ggg 224
Thr Leu Ala Thr Thr Ser Lys Asn Ile Thr Ser Gly Leu His Phe Gly
10 15 20

ctt gtg aac atc tct ggc aac aat gag tct acc ttg aac tgt tca cag 272
Leu Val Asn Ile Ser Gly Asn Asn Glu Ser Thr Leu Asn Cys Ser Gln
25 30 35

aaa cca tca gat aag cat tta gat gca att cct att ctt tac tac att 320
Lys Pro Ser Asp Lys His Leu Asp Ala Ile Pro Ile Leu Tyr Tyr Ile
40 45 50

ata ttt gta att gga ttt ctg gtc aat att gtc gtg gtt aca ctg ttt 368
Ile Phe Val Ile Gly Phe Leu Val Asn Ile Val Val Val Thr Leu Phe
55 60 65

tgt tgt caa aag ggt cct aaa aag gtt tct agc ata tac atc ttc aac 416
Cys Cys Gln Lys Gly Pro Lys Lys Val Ser Ser Ile Tyr Ile Phe Asn
70 75 80 85

ctc gct gtg gct gat tta ctc ctt ttg gct act ctt cct cta tgg gca 464
Leu Ala Val Ala Asp Leu Leu Leu Leu Ala Thr Leu Pro Leu Trp Ala
90 95 100

acc tat tat tct tat aga tat gac tgg ctc ttt gga cct gtg atg tgc 512
Thr Tyr Tyr Ser Tyr Arg Tyr Asp Trp Leu Phe Gly Pro Val Met Cys
105 110 115

aaa gtt ttt ggt tct ttt ctt acc ctg aac atg ttt gca agc att ttt 560
Lys Val Phe Gly Ser Phe Leu Thr Leu Asn Met Phe Ala Ser Ile Phe
120 125 130

ttt atc acc tgc atg agt gtt gat agg tac caa tct gtc atc tac ccc 608
Phe Ile Thr Cys Met Ser Val Asp Arg Tyr Gln Ser Val Ile Tyr Pro
135 140 145

ttt ctg tct caa aga aga aat ccc tgg caa gca tct tat ata gtt ccc 656

```
Phe Leu Ser Gln Arg Arg Asn Pro Trp Gln Ala Ser Tyr Ile Val Pro
150                 155                 160                 165

ctt gtt tgg tgt atg gcc tgt ttg tcc tca ttg cca aca ttt tat ttt      704
Leu Val Trp Cys Met Ala Cys Leu Ser Ser Leu Pro Thr Phe Tyr Phe
                170                 175                 180

cga gac gtc aga acc att gaa tac tta gga gtg aat gct tgc att atg      752
Arg Asp Val Arg Thr Ile Glu Tyr Leu Gly Val Asn Ala Cys Ile Met
            185                 190                 195

gct ttc cca cct gag aaa tat gcc caa tgg tca gct ggg att gcc tta      800
Ala Phe Pro Pro Glu Lys Tyr Ala Gln Trp Ser Ala Gly Ile Ala Leu
        200                 205                 210

atg aaa aat atc ctt ggt ttt att atc cct tta ata ttc ata gca aca      848
Met Lys Asn Ile Leu Gly Phe Ile Ile Pro Leu Ile Phe Ile Ala Thr
    215                 220                 225

tgc tat ttt gga att aga aaa cac tta ctg aag acg aat agc tat ggg      896
Cys Tyr Phe Gly Ile Arg Lys His Leu Leu Lys Thr Asn Ser Tyr Gly
230                 235                 240                 245

aag aac agg ata acc cgt gac caa gtc ctg aag atg gca gct gct gtt      944
Lys Asn Arg Ile Thr Arg Asp Gln Val Leu Lys Met Ala Ala Ala Val
                250                 255                 260

gtt ctg gcc ttc atc att tgg tgc ctt ccc ttc cat gtt ctg acc ttc      992
Val Leu Ala Phe Ile Ile Trp Cys Leu Pro Phe His Val Leu Thr Phe
            265                 270                 275

ctg gat gct ctg gcc tgg atg ggt gtc att aat agc tgc gaa gtt ata     1040
Leu Asp Ala Leu Ala Trp Met Gly Val Ile Asn Ser Cys Glu Val Ile
        280                 285                 290

gca gtc att gac ctg gca ctt cct ttt gcc atc ctc ttg gga ttc acc     1088
Ala Val Ile Asp Leu Ala Leu Pro Phe Ala Ile Leu Leu Gly Phe Thr
    295                 300                 305

aac agc tgc gtt aat ccg ttt ctg tat tgt ttt gtt gga aac cgg ttc     1136
Asn Ser Cys Val Asn Pro Phe Leu Tyr Cys Phe Val Gly Asn Arg Phe
310                 315                 320                 325

caa cag aag ctc cgc agt gtg ttt agg gtt cca att act tgg ctc caa     1184
Gln Gln Lys Leu Arg Ser Val Phe Arg Val Pro Ile Thr Trp Leu Gln
                330                 335                 340

ggg aaa aga gag agt atg tct tgc cgg aaa agc agt tct ctt aga gaa     1232
Gly Lys Arg Glu Ser Met Ser Cys Arg Lys Ser Ser Ser Leu Arg Glu
            345                 350                 355

atg gag acc ttt gtg tct taa acggagagca aaatgcatgt aatcaacatg        1283
Met Glu Thr Phe Val Ser  *
        360

gctacttgct ttgaggctca ccagaattat ttttaagtgg ttttaataaa ataataaaat   1343
ttcccctaat cttttctgaa tcttctgaaa ccaaatgtaa ctatgtttat cgtccagtga   1403
ctttcaggaa tgcccattgt tttctgatat gtttgtacaa gatttcattg gtgagacata   1463
```

tttacaacct agaagtaact ggtgatatat ctcaaattgt aattaataat agattgtgaa 1523
taatgatttg gggattcaga tttctctttg aaacatgctt gtgtttctta gtggggtttt 1583
atatccattt ttatcaggat ttcctcttga accagaacca gtctttcaac tcattgcatc 1643
atttacaaga caacattgta agagagatga gcacttctaa gttgagtata ttataataga 1703
ttagtactgg attattcagg ctttaggcat atgcttcttt aaaaacgcta taaattatat 1763
tcctcttgca tttcacttga gtggaggttt atagttaatc tataactaca tattgaatag 1823
ggctaggaat atagattaaa tcatactcct atgctttagc ttatttttac agttatagaa 1883
agcaagatgt actataacat agaattgcaa tctataatat ttgtgtgttc actaaactct 1943
gaataagcac tttttaaaaa actttctact cattttaatg attgtttaaa ggtttctatt 2003
ttctctgata ctttttttgaa atcagtaaac actgtgtatt gttgtaaaat gtaaaggtca 2063
cttttcacat ccttgacttt ttagatgtgc tgctttgata tataggacat tgatttgatt 2123
tttattatta atgctttggt tctgggttgt ttcctaaaat atctgggtgg cttaaaaaaa 2183
actctttaac ttgtaataaa cccttaactg gcataggaaa tggtatccag aatggaattt 2243
tgctacatgg ggtctgggtg ggggcaaaga gacccagtca attacatgtt tggtaccaag 2303
aaaggaacct gtcagggcag tacaatgtga ctttgaaaat atataccgtg gggggtagttt 2363
taccctatat ctataaacac tgtttgttcc agaatctgta tgattctatg gagctatttt 2423
aaaccaattg caggtctaga 2443

<210> 40
<211> 363
<212> PRT
<213> Homo sapien

<400> 40
Met Lys Gly Asn Ser Thr Leu Ala Thr Thr Ser Lys Asn Ile Thr Ser
1 5 10 15
Gly Leu His Phe Gly Leu Val Asn Ile Ser Gly Asn Asn Glu Ser Thr
20 25 30
Leu Asn Cys Ser Gln Lys Pro Ser Asp Lys His Leu Asp Ala Ile Pro
35 40 45
Ile Leu Tyr Tyr Ile Ile Phe Val Ile Gly Phe Leu Val Asn Ile Val
50 55 60
Val Val Thr Leu Phe Cys Cys Gln Lys Gly Pro Lys Lys Val Ser Ser
65 70 75 80
Ile Tyr Ile Phe Asn Leu Ala Val Ala Asp Leu Leu Leu Leu Ala Thr
85 90 95
Leu Pro Leu Trp Ala Thr Tyr Tyr Ser Tyr Arg Tyr Asp Trp Leu Phe
100 105 110
Gly Pro Val Met Cys Lys Val Phe Gly Ser Phe Leu Thr Leu Asn Met
115 120 125
Phe Ala Ser Ile Phe Phe Ile Thr Cys Met Ser Val Asp Arg Tyr Gln
130 135 140
Ser Val Ile Tyr Pro Phe Leu Ser Gln Arg Arg Asn Pro Trp Gln Ala
145 150 155 160
Ser Tyr Ile Val Pro Leu Val Trp Cys Met Ala Cys Leu Ser Ser Leu
165 170 175
Pro Thr Phe Tyr Phe Arg Asp Val Arg Thr Ile Glu Tyr Leu Gly Val
180 185 190
Asn Ala Cys Ile Met Ala Phe Pro Pro Glu Lys Tyr Ala Gln Trp Ser
195 200 205
Ala Gly Ile Ala Leu Met Lys Asn Ile Leu Gly Phe Ile Ile Pro Leu
210 215 220
Ile Phe Ile Ala Thr Cys Tyr Phe Gly Ile Arg Lys His Leu Leu Lys
225 230 235 240
Thr Asn Ser Tyr Gly Lys Asn Arg Ile Thr Arg Asp Gln Val Leu Lys
245 250 255
Met Ala Ala Ala Val Val Leu Ala Phe Ile Ile Trp Cys Leu Pro Phe
260 265 270

```
His Val Leu Thr Phe Leu Asp Ala Leu Ala Trp Met Gly Val Ile Asn
        275                 280                 285
Ser Cys Glu Val Ile Ala Val Ile Asp Leu Ala Leu Pro Phe Ala Ile
    290                 295                 300
Leu Leu Gly Phe Thr Asn Ser Cys Val Asn Pro Phe Leu Tyr Cys Phe
305                 310                 315                 320
Val Gly Asn Arg Phe Gln Gln Lys Leu Arg Ser Val Phe Arg Val Pro
                325                 330                 335
Ile Thr Trp Leu Gln Gly Lys Arg Glu Ser Met Ser Cys Arg Lys Ser
            340                 345                 350
Ser Ser Leu Arg Glu Met Glu Thr Phe Val Ser
        355                 360
```

<210> 41
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 41
actgcctgat aaccatgctg 20

<210> 42
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 42
atacttacac accaggaggg 20

<210> 43
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 43
atgcctgctc caaaggcac 19

<210> 44
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 44
atgcctgctc caaaggcacc 20

<210> 45

<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 45
atgcctgctc caaaggcaca t 21

<210> 46
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 46
tacttctggt tctctgagcg 20

<210> 47
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 47
actcaccttg aactcgtctc 20

<210> 48
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 48
tggttctctg agcgagtctt 20

<210> 49
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 49
tggttctctg agcgagtctt c 21

<210> 50
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 50
tggttctctg agcgagtctt tc 22

<210> 51
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 51
tgcagatgga ctttggcttc 20

<210> 52
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 52
tgcttgcctt ctgctacaag 20

<210> 53
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 53
cttccctgag cacctgctg 19

<210> 54
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 54
cttccctgag cacctgctgg t 21

<210> 55
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 55
cttccctgag cacctgctga 20

<210> 56
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 56
aacagctcag gacgaaactg 20

<210> 57
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 57
agaaggagtt gaccttgtcc 20

<210> 58
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 58
ggaagctcaa gtggccttc 19

<210> 59
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 59
ggaagctcaa gtggccttcc 20

<210> 60
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 60
ggaagctcaa gtggccttca ac 22

<210> 61
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 61
aagtcactgg cagagctgg 19

<210> 62
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 62
gcaccagggc tttgttgaag 20

<210> 63
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 63
ttttccccgt agggctcca 19

<210> 64
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 64
ttttccccgt agggctccac 20

<210> 65
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 65
ttttccccgt agggctccag c 21

<210> 66
<211> 20
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer

<400> 66
tgcagaagtc actggcagag 20

<210> 67
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 67
gttgaagttt tccccgtagg 20

<210> 68
<211> 19
<212> DNA
<213> Aritificial sequence

<400> 68
actcctccac ctgctggtc 19

<210> 69
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 69
actcctccac ctgctggtcc 20

<210> 70
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 70
actcctccac ctgctggtct a 21

<210> 71
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 71

aggacgtgcg tggcaacctg 20

<210> 72
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 72
agctctgcca gtgacttctg 20

<210> 73
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 73
gtgacttctg cagcccctc 19

<210> 74
<211> 20
<212> DNA
<213> Artitifical sequence

<400> 74
gtgacttctg cagcccctca 20

<210> 75
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 75
gtgacttctg cagcccctcg gt 22

<210> 76
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 76
cctgaccttc cagatgaag 19

<210> 77
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 77
tcaggttgcc acgcacgtc 19

<210> 78
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 78
caggatctcg gccagtgc 18

<210> 79
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 79
caggatctcg gccagtgcc 19

<210> 80
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 80
caggatctcg gccagtgctg 20

<210> 81
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 81
acctgcgaga gcttcagcag 20

<210> 82
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 82
tctccatgcg ctgtgcgtag 20

<210> 83
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 83
agctgcgcac ccaggtca 18

<210> 84
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 84
agctgcgcac ccaggtcaa 19

<210> 85
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 85
agctgcgcac ccaggtcagc 20

<210> 86
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 86
tgtccaagga gctgcaggc 19

<210> 87
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 87
cttacgcagc ttgcgcaggt 20

<210> 88
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 88
gcggacatgg aggacgtg 18

<210> 89
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 89
gcggacatgg aggacgtgc 19

<210> 90
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 90
gcggacatgg aggacgtgtg 20

<210> 91
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 91
gttgtagaaa gaaccgctgc 20

<210> 92
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 92
gagaacgagt cttcaggtac 20

<210> 93
<211> 21
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer

<400> 93
acaatctggg ctatgagatc a 21

<210> 94
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 94
acaatctggg ctatgagatc aa 22

<210> 95
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 95
acaatctggg ctatgagatc agt 23

<210> 96
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 96
cactctacac tgcatgtctc 20

<210> 97
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 97
acccttctga aaaggagagg 20

<210> 98
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Primer

<400> 98
gaggagagac aaggcagata 20

<210> 99
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 99
gaggagagac aaggcagata t 21

<210> 100
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 100
gaggagagac aaggcagata gt 22

<210> 101
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 101
aaaggttcag ttgctgctgc 20

<210> 102
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 102
gctggggaag gtctaataac 20

<210> 103
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 103

gttgctgctg cctcgaatc 19

<210> 104
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 104
gttgctgctg cctcgaatcc 20

<210> 105
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 105
gttgctgctg cctcgaatct g 21

<210> 106
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 106
cgtctttctc cagatgatgc 20

<210> 107
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 107
agtgtcctat gggctgtttg 20

<210> 108
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 108
ggatgccatt catacctta c 21

<210> 109

<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 109
ggatgccatt cataccttta cc 22

<210> 110
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 110
ggatgccatt cataccttta cgc 23

<210> 111
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 111
tgggaaaaca gtgcagtgtg 20

<210> 112
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 112
tgatcgtctt cagaacgagg 20

<210> 113
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 113
ccagaccatc atcatcccat gga 23

<210> 114
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 114
ccagaccatc atcccatgga a 21

<210> 115
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 115
ccagaccatc atcccatgga gc 22

<210> 116
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 116
cagcaatcgt ctttctccag 20

<210> 117
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 117
tcctatgggc tgtttgatgc 20

<210> 118
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 118
gtctttctcc agatgatgcc a 21

<210> 119
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 119
gtctttctcc agatgatgcc aa 22

<210> 120
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 120
gtctttctcc agatgatgcc agt 23

<210> 121
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 121
agcggataac aatttcacac agg 23

<210> 122
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 122
ggcgcacgcc tccacg 16